

UNIVERSITY OF NOTTINGHAM

DEPARTMENT OF CIVIL ENGINEERING

DETERMINATION OF SATELLITE ORBITS AND

THE GLOBAL POSITIONING SYSTEM

by

Loukis George Agrotis, B.Sc., M.Inst.C.E.S.

Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy

October 1984

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	ix
ACKNOWLEDGEMENTS	xi
LIST OF FIGURES	xiii
 CHAPTER 1	
INTRODUCTION	1
 CHAPTER 2	
SATELLITE ORBIT DETERMINATION	
 2.1	
Basic Concepts	7
 2.2	
COORDINATE REFERENCE FRAMES AND TIME SCALES	
 2.2.1	
Reference Frames and Time	10
 2.2.2	
Precession and Nutation	13
 2.2.3	
Earth Rotation and Polar Motion	21
 2.2.4	
Coordinate Transformations	28
 2.3	
FORCE MODEL COMPONENTS	
 2.3.1	
Introduction	28
 2.3.2	
Earth Gravitational Attraction	29
 2.3.3	
Moon, Sun and Planetary Attractions	32
 2.3.4	
Solid Earth Tides	35
 2.3.5	
Ocean Tides	43
 2.3.6	
Empirical Accelerations	45

		<u>Page</u>
2.3.7	Air Drag	46
2.3.8	Solar and Albedo Radiation	49
2.3.9	Satellite Thrust	55
2.4	NUMERICAL INTEGRATION OF THE EQUATIONS OF MOTION	
2.4.1	Equations of Motion	56
2.4.2	Numerical Integration of Differential Equations	57
2.5	LEAST SQUARES ADJUSTMENT AND PARTIAL DERIVATIVES	
2.5.1	The Satellite Observations	65
2.5.2	Least Squares Adjustment	67
2.5.3	The Orbit Determination Observation Equations	73
2.5.4	Orbit Determination Adjustment Requirements	82
CHAPTER 3	UNIVERSITY OF NOTTINGHAM ORBIT DETERMINATION SOFTWARE	
3.1	Introduction	86
3.2	CHEBYSHEV POLYNOMIAL PROGRAM (CHEBPOL)	
3.2.1	Data Input	90
3.2.2	Program Description and Output	90

3.3	THE SATELLITE ORBIT INTEGRATION PROGRAM (ORBIT)	
3.3.1	Numerical Integration and Force Model Definition	94
3.3.2	Input Requirements	100
3.3.3	General Outline	104
3.3.4	Program Output	111
3.3.5	Software Debugging	112
3.4	THE SATELLITE ORBIT ANALYSIS PROGRAM (SOAP)	
3.4.1	Program Input and Output	114
3.4.2	General Description	118
CHAPTER 4	ANALYSIS OF SATELLITE TRACKING DATA AND APPLICATIONS OF ORBIT DETERMINATION	
4.1	Introduction	122
4.2	RESULTS OF ANALYSIS	
4.2.1	LAGEOS Tracking Data	123
4.2.2	Effect of Force Model Parameters and Geopotential Models	127
4.2.3	Solutions for Tracking Station Coordinates and Satellite Starting Elements	138
4.2.4	Polar Motion Solutions	143

		<u>Page</u>
4.2.5	Effect of Earth Rotation	149
4.3	Applications of Precise Orbit Determination	153
CHAPTER 5	THE GLOBAL POSITIONING SYSTEM	
5.1	Introduction	158
5.2	GENERAL SYSTEM DESCRIPTION	
5.2.1	Space Segment	160
5.2.2	Control Segment	163
5.2.3	User Segment	166
5.3	GPS SATELLITE SIGNALS	
5.3.1	Signal Structure and Receiver Measurement Sequence	167
5.3.2	The C/A Code	172
5.3.3	The P Code	176
5.3.4	Satellite Data Message	177
5.4	MODES OF OBSERVATION AND ADJUSTMENT FOR PRECISE POSITIONING	
5.4.1	Instantaneous Navigation Principles	180
5.4.2	Pseudo-Range Measurements	185
5.4.3	Phase (Doppler) Observations	189
5.4.4	Interferometric Techniques	196

		<u>Page</u>
5.4.5	Reduction of Observations	204
5.4.5.1	Ionospheric Corrections	205
5.4.5.2	Tropospheric Corrections	209
5.4.5.3	Relativity	212
5.5	GPS Geodetic User Equipment	212
CHAPTER 6	GPS SOFTWARE DEVELOPMENT AND ANALYSIS OF SAMPLE DATA	
6.1	Introduction	217
6.2	THE GPS DATA	
6.2.1	The NAVSTAR Geodetic Receiver System (NGRS)	218
6.2.2	Data Specification	219
6.3	NGRS DATA PRE-PROCESSING SOFTWARE	
6.3.1	Aims of Pre-Processing	224
6.3.2	General Outline	226
6.3.2.1	Smoothed Pseudo-Range Mode	227
6.3.2.2	Biased Range (Phase) Mode	230
6.4	GPS LEAST SQUARES ADJUSTMENT PROGRAM (GPSPROG)	
6.4.1	Introduction and Data Input	237
6.4.2	Program Description	239

		<u>Page</u>
6.4.3	Future Development	245
6.5	ANALYSIS OF THE NGRS-2 DATA	
6.5.1	Introduction	248
6.5.2	Biased Range Solutions	250
6.5.3	Smoothed Pseudo-Range Solutions	264
6.5.4	Single-Frequency Solutions	269
CHAPTER 7	CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK	
7.1	CONCLUSIONS	
7.1.1	Conclusions on Orbit Determination	275
7.1.2	Conclusions on the Global Positioning System	276
7.2	SUGGESTIONS FOR FURTHER WORK	
7.2.1	Suggestions for Further Work on Orbit Determination	278
7.2.2	Suggestions for Further Work on GPS	279
APPENDICES		
A	ROTATION AND REFLECTION MATRICES	281
B	COORDINATE REPRESENTATIONS	283
C	ORBIT DETERMINATION PARTIAL DERIVATIVES	291

	<u>Page</u>
D METHODS OF NUMERICAL INTEGRATION FOR ORBITS	301
E POLYNOMIAL REPRESENTATION AND INTERPOLATION	307
F WEIGHTED LEAST SQUARES	310
G COMPUTATION OF GPS TIME AND SATELLITE COORDINATES	317
REFERENCES AND BIBLIOGRAPHY	331

DETERMINATION OF SATELLITE ORBITS AND THE GLOBAL
POSITIONING SYSTEM

ABSTRACT

An artificial satellite orbit determination (OD) computer program is the most essential tool in satellite geodesy. Such a program has been developed at Nottingham as part of this research and was tested with Satellite Laser Ranging (SLR) observations of the Laser Geodynamics Satellite (LAGEOS).

This thesis describes the basic theory behind orbit determination and the software development at Nottingham. It includes details of the adopted force model, coordinate reference frames, and numerical integration and interpolation techniques. It is also explained how several geodetic parameters can be determined. The thesis discusses the results of two separate determinations of the LAGEOS orbit with an emphasis on the solutions for station coordinates and for earth rotation and polar motion.

The NAVSTAR Global Positioning System (GPS) is on schedule to replace Transit as the most important satellite navigation system. When fully operational, in 1988, it will consist of 18 satellites which will provide continuous global coverage. This thesis describes the Global Positioning System and outlines the theory behind the most accurate techniques of adjustment of the GPS observables. It derives the equations for interferometric techniques and shows that, by differencing the observations, several undesirable unknowns can be eliminated.

GPS data from the NAVSTAR Geodetic Receiver System (NGRS) have been provided for Nottingham by the US Defence Mapping Agency (DMA). The thesis describes the software development to analyse these data and gives the results of several solution schemes to derive the absolute coordinates of the NGRS antenna. It is also shown how the software can be modified to incorporate interferometric techniques. Significant improvements over the NGRS solutions can be expected when GPS is fully operational, with refinements in both receiver hardware and software.

ACKNOWLEDGEMENTS

This thesis is the result of three years of fulfilling research in the Department of Civil Engineering at the University of Nottingham. The work was carried out with the support of the Heads of Department, Professor R. C. Coates and Professor P. S. Pell, and was funded by a University of Nottingham Postgraduate Studentship and by an Overseas Research Students Fees Support Scheme Award. The author was also sponsored by British Petroleum.

The author is indebted to his supervisor, Professor V. Ashkenazi, for his guidance, support, and encouragement throughout this study.

The data and other documentation vital for this research have been provided by the Royal Greenwich Observatory and by the US Defence Mapping Agency. Sincere thanks are extended to the staff of both these establishments and especially to Dr. A. T. Sinclair, Dr. G. A. Wilkins, Mr. G. Appleby, Mr. M. M. Macomber, Mr. K. I. Dougherty and Dr. M. Kumar. Other scientists who have given valuable advice include Dr. P. J. Hargrave, Dr. C. C. Goad, Dr. R. J. Anderle, Dr. B. R. Hermann, Dr. G. Lachapelle, Dr. C. C. Counselman, Dr. P. F. MacDoran and Mr. J. Dow.

The author acknowledges the assistance received from the other members of the Nottingham Surveying Group, notably from Dr. R. Wood, who gave advice on software development, and from Dr. A. H. Dodson, Mr. T. Moore, Mr. P. Howard and Dr. S. Grist.

The author expresses his gratitude to his parents for their encouragement and support and to Libby, whose patience and help have been invaluable throughout the course of this work.

Finally the author wishes to thank Mrs. Beryl Greaves for typing this thesis in her neat and efficient style.

LIST OF FIGURES

<u>Figure No</u>	<u>Title</u>	<u>Page</u>
2.I	Orbit Determination Block Diagram	8
2.II	The Celestial Sphere and Precession	15
2.III	The Fundamental Arguments	18
2.IV	The Nutation Angles	19
2.V	Earth Rotation and Polar Motion	23
2.VI	Coordinate Reference Frames and Time Scales	27
2.VII	Third Body Gravitational Attraction	33
2.VIII	Earth Tides	37
2.IX	Satellite Velocity Relative to Atmosphere	47
2.X	Umbra and Penumbra	52
2.XI	Shadow Test for Satellite	53
3.I	Nottingham Orbit Determination Software	87
3.II	Flow Diagram for CHEBPOL	93
3.III	Typical Magnitudes of LAGEOS Force Model Constituents	98
3.IV	Data Input and Output for ORBIT	101
3.V	ORBIT Flow Chart	108
3.VI	SOAP Input and Output	115
4.I	SL5 Coordinates of LAGEOS Tracking Stations	124
4.II	LAGEOS Tracking Sites	125
4.III	Distribution of Normal Points for the Two Data Sets	126

<u>Figure No</u>	<u>Title</u>	<u>Page</u>
4.IV	Range Residuals From Solution for Satellite State Vector and Station Coordinates	130
4.V	Range Residuals From Solution for Satellite State Vector, Station Coordinates, and GM	131
4.VI	Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a	132
4.VII	Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a	133
4.VIII	Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a	134
4.IX	Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a	135
4.X	Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a	136
4.XI	Force Model Parameter Solution	137
4.XII	Solutions for Tracking Station Coordinates	140
4.XIII	Comparison Between DS1, DS2, and SL5 Station Coordinates	141
4.XIV	Solutions for LAGEOS Starting Elements	142
4.XV	Polar Motion Solutions	144

<u>Figure No</u>	<u>Title</u>	<u>Page</u>
4.XVI	Comparison Between DS1, DS2, and SL5 Station Coordinates After Polar Motion Solution	146
4.XVII	Daily Earth Rotation Solutions	151
4.XVIII	Comparison of Daily Changes in l.o.d. Derived from Different Sources	152
5.I	Operational GPS Satellite Configuration	161
5.II	ICS and OCS Monitor Stations	165
5.III	PRN Modulations	170
5.IV	L1 Signal Power Spectral Density	170
5.V	4-Stage Linear Feedback Shift Register	174
5.VI	Satellite Interferometry	197
6.I	NGRS-2 Observation Schedule	220
6.II	NGRS-2 60-Second Phase Measurement	222
6.III	Block Diagram of GPS Software	225
6.IV	Statistics of NGRS-2 Pre-Processed Observations	249
6.V	Data Set 1: Pass-By-Pass Biased Range Solutions	253
6.VI	Data Set 2: Pass-By-Pass Biased Range Solutions	254
6.VII	Data Set 3: Pass-By-Pass Biased Range Solutions	255
6.VIII	Data Set 1: Pass-By-Pass Biased Range Solutions	258
6.IX	Data Set 2: Pass-By-Pass Biased Range Solutions	259
6.X	Data Set 3: Pass-By-Pass Biased Range Solutions	260
6.XI	Biased Range Final Solutions	263

<u>Figure No</u>	<u>Title</u>	<u>Page</u>
6.XII	Data Set 1: Pass-By-Pass Smoothed Pseudo-Range Solutions	265
6.XIII	Data Set 1: Pass-By-Pass Smoothed Pseudo-Range Solutions	266
6.XIV	Smoothed Pseudo-Range Final Solutions	268
6.XV	L1 Biased Range Solutions	270
6.XVI	L1 Smoothed Pseudo-Range Solutions	271
6.XVII	Comparison Between Biased Range and Smoothed Pseudo-Range Solutions	273

CHAPTER ONE

INTRODUCTION

1. INTRODUCTION

Geodesy has been defined by various people (Moritz, 1984) as the science of measuring and mapping the surface and gravity field of the earth. Traditionally, the techniques available to the geodesist comprised terrestrial angle and distance measurements as well as astro-geodetic and gravimetric methods. In 1957 George C. Weiffenbach and William H. Guier of the Applied Physics Laboratory, John Hopkins University, showed that they could determine the orbit of the first Sputnik by making measurements of the Doppler shift of the received signals from this spacecraft. From this initial concept evolved the science of satellite geodesy which, in the last decade, has shown sub-metre absolute positioning and sub-decimetre relative positioning capabilities. Satellite techniques have also resulted in global determinations of the earth's gravity field.

A large number of artificial satellites have been launched over the last 25 years. Several of these have been used for geodetic applications. The most notable are the Laser Geodynamics Satellite (LAGEOS) and the satellites of the Navy Navigation Satellite System (NNSS or Transit) and of the NAVSTAR Global Positioning System (GPS). Each satellite is tracked by observations from a network of globally distributed tracking stations. The observations are then incorporated in an orbit determination (OD) process in order to compute the satellite orbit and a number of geodetic parameters.

Orbit determination software is thus the basic tool in satellite geodesy. It utilizes the tracking observations to derive a number of unknown quantities. These can be the satellite orbital elements, the locations of the tracking stations, earth rotation

and polar motion values, tidal coefficients, and geopotential coefficients. The choice of unknowns depends on the objectives of the particular satellite mission and on the capabilities of the tracking network. The most well known OD computer programs are the CELEST program (O'Toole, 1976), which is used in the Transit system, and the Geodyn program (Martin et al, 1980), which is commonly employed with Satellite Laser Ranging (SLR) observations.

The heart of an OD process is a force model which defines the forces acting on the satellite. The model consists of constituents arising from the gravitational attractions of the earth, moon, sun, and planets, and from surface forces such as solar radiation pressure and air drag. The vector sum of all these constituents gives the resultant acceleration of the satellite as a function of its position and velocity. The acceleration is then numerically integrated twice to give the velocity and position of the satellite as a function of time. The numerical integration is started by assuming some initial values of the satellite position and velocity, the satellite starting elements or state vector. These elements need only be approximate, since they can be determined by least squares from the tracking data. The least squares adjustment also determines a number of other unknowns, as mentioned above.

Satellite Laser Range (SLR) observations are the most accurate tracking measurements available. The present third generation SLR facilities are capable of ranging accuracies of 3-5 cm over distances of the order of 7000 Km. Radio frequency (RF) observations are less accurate and are affected by errors arising from ionospheric delays. However, since SLR facilities are expensive and are limited by atmospheric conditions (clouds), RF tracking is the most commonly

used technique. Three types of RF measurements are possible :
the first type involve the measurements of the Doppler shift of the received signals. The second type are range measurements similar to SLR. The last type are termed pseudo-ranges and these are range measurements which include a clock bias between receiver and satellite clocks.

Orbit determination software has been developed at Nottingham as part of this research. Since SLR is the most important technique for precise geodetic applications, the software was written primarily for use with range observations to LAGEOS. LAGEOS tracking data are being provided for Nottingham by the Royal Greenwich Observatory and have been used to test the software and to establish the capabilities of OD. Special emphasis has been given to the use of the SLR data to determine earth rotation and polar motion, in view of the present international project to Monitor Earth Rotation and to Intercompare the Techniques of observation and adjustment (MERIT). The accuracy and repeatability of the tracking station coordinates has also been investigated.

A most important application of satellite orbit determination is in quasi-instantaneous positioning (navigation) and in the computation of precise absolute or relative coordinates on the surface of the earth (surveying and geodesy). The very first satellite system designed primarily for positioning was the Transit system . The first Transit satellite was launched by the US Navy in 1960, and the system was made available to civilian users in 1967. The basic principle of operation is for the satellite orbits to be determined from observations by a network of tracking stations.

The complex and computationally expensive orbit determinations are performed by a central processing facility. The users of Transit then employ this orbital information and observations collected by their receiving equipment to determine their own positions. A set of predicted orbital elements, the Broadcast Ephemeris, based on the most recent tracking data, is transmitted by the Transit satellites. A more precise post-mission ephemeris, the Precise Ephemeris, is also available to certain bona fide users.

In the twenty or more years that Transit has been in continuous operation it has performed beyond all expectations. Three of the four operational satellites were built from a 1963 design and are still functioning after 10-17 years in service. Transit observations result in absolute positioning accuracies of 2-5 m after time spans of 2-3 days. Specialized methods of adjustment (Gough, 1978 and Sykes, 1979) yield relative positions to sub-metre accuracies over baseline lengths of hundreds of kilometres. Transit, however, is incapable of instantaneous navigation and, to satisfy US military requirements, it is to be succeeded by a new system, the NAVSTAR Global Positioning System (GPS).

GPS is presently in its Full Scale Engineering Development Phase, with five operational satellites. When fully operational, in 1988, the system will consist of 18 satellites in circular orbits at inclinations of 55° to the equator. These will travel at altitudes of approximately 20000 Km and will orbit the earth every 12 hours. A minimum of 4 satellites will always be in view anywhere on the earth.

The operational details of GPS are very similar to Transit. The satellites are tracked by a network of Monitor Stations (MS's). A master Control Station (MCS) computes the satellite orbits, based on the most recent observations, and uploads these and other relevant information to the satellites via a number of Ground Antennas (GA's). The satellites, in turn, transmit their orbital elements to the GPS users as part of the satellite navigation message.

The basic GPS observable is a pseudo-range, which is derived by making use of coded ranging (timing) information modulated on the satellite signals. For an instantaneous solution, a navigation receiver has to make at least four simultaneous pseudo-range measurements to different satellites. These enable the receiver processor to solve for the cartesian coordinates of the antenna and for the receiver clock bias from GPS time.

Geodetic applications have no requirement for an instantaneous capability. For this reason, techniques have been developed which post-process the GPS observations to derive precise absolute coordinates. Interferometric techniques utilize simultaneous observations from two or more receiver sites to compute even more precise relative coordinates.

The first GPS geodetic receiver was the NAVSTAR Geodetic Receiver System (NGRS), developed by the Naval Surface Weapons Center (NSWC). It uses an atomic (cesium) clock but can only observe one satellite at a time. NGRS data covering a time span of 10 days have been supplied for Nottingham University by the US Defence Mapping Agency (DMA). These have been used to develop and test GPS absolute positioning software. The software was written with the intention of incorporating interferometric techniques after slight modifications.

The NGRS is limited by its inability to track more than one satellite at a time. It has not therefore produced accuracies which exceed those attainable from the Transit system.

The present GPS geodetic receiver development is concentrated on instruments capable of observing a number of satellites simultaneously. These receivers are the ones which are expected to show a significant improvement in accuracy over Transit. One such receiver system, the Macrometer, can operate without knowledge of the GPS ranging codes. This feature is important since access to the codes may be denied to non-military users.

This thesis covers two major topics; satellite orbit determination, and the Global Positioning System. Chapter 2 describes the basic theory behind the orbit determination process. The Nottingham OD software is described in Chapter 3. Chapter 4 discusses the solutions for various geodetic quantities from two determinations of the LAGEOS orbit. The details of the Global Positioning System and the theory behind the techniques of observation and adjustment are given in Chapter 5. Chapter 6 describes the GPS software developed at Nottingham and discusses the results from the NGRS data. The thesis is concluded in Chapter 7.

CHAPTER TWO

SATELLITE ORBIT DETERMINATION

2.1 Basic Concepts

Satellite orbit determination (OD) is the process through which a set of satellite tracking data, observed by a number of widely distributed tracking stations, is used to determine a precise satellite ephemeris. Several other parameters can also be derived. These include tracking station coordinates, values of polar motion and earth rotation, geopotential (spherical harmonic) coefficients and Love numbers. The flow chart in fig.2.I illustrates the basic principles involved in OD. A force model is used to accurately describe the various forces acting on the satellite, which consists of components arising from gravitational, surface and other forces. The vector sum of the separate components gives the resultant force and hence the resultant acceleration of the satellite. This is a function of satellite position and, where drag is involved, of velocity. The acceleration is numerically integrated, once to obtain velocity and twice to obtain position, as a function of time. The integration is extended to sufficient steps to cover the observational time span. For the process to begin, a set of starting elements comprising the satellite position and velocity at an initial epoch, t_0 , is required. This satellite state vector does not need to be precisely known. The observations taken by the network of tracking stations are used to obtain improved values of the state vector, in a least squares solution which can also determine various force model parameters, as well as tracking station coordinates. If the corrections to the provisional values of the various unknowns are large, the orbit determination is repeated until the values of the corrections become negligible.

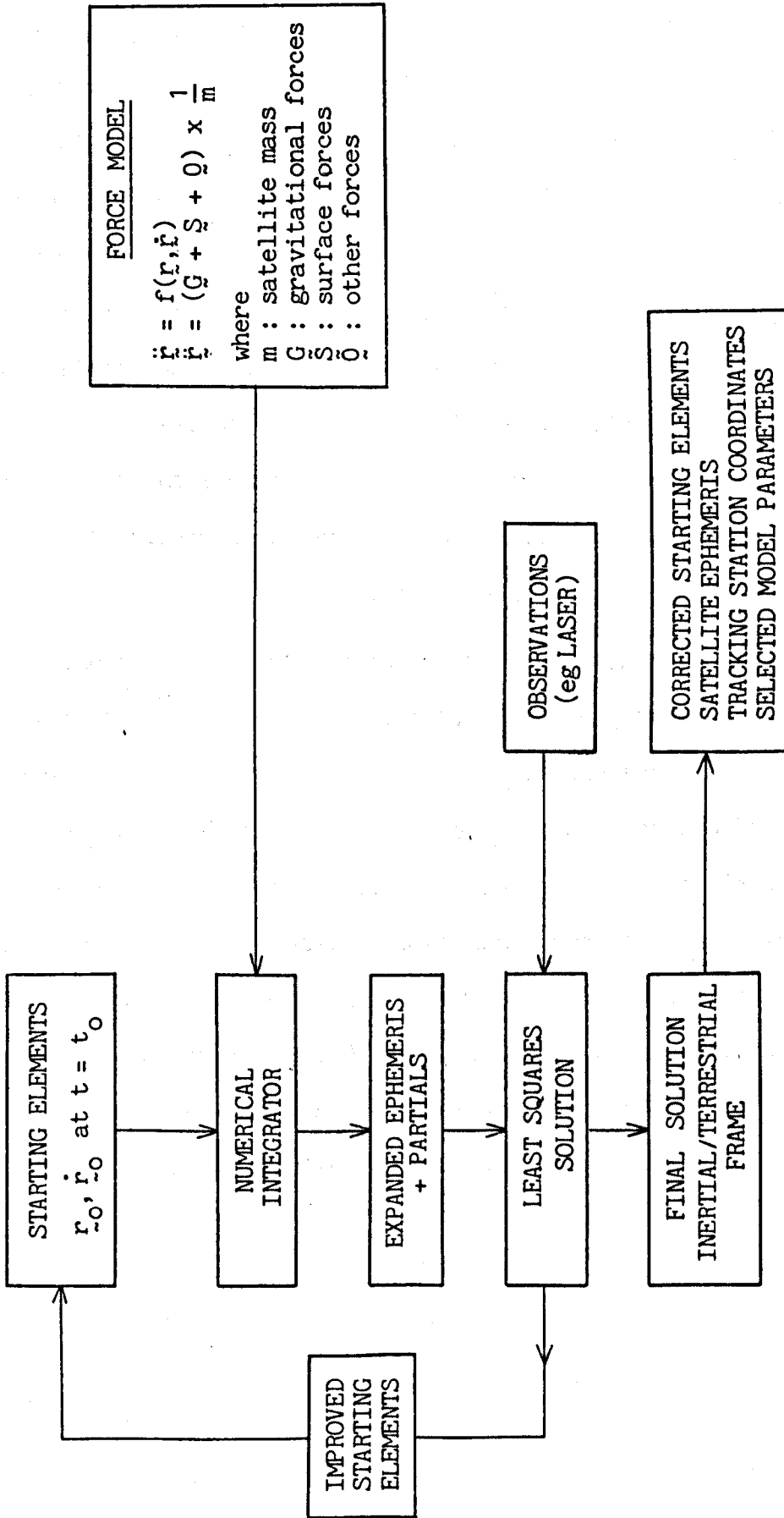


Fig. 2.I Orbit Determination Block Diagram

The choice of unknowns depends on the objectives of the OD process. When the satellite ephemeris is the only requirement, as is the case for the US Navy Navigation Satellite System (NNSS or Transit) (Anderle, 1974) and for the NAVSTAR Global Positioning System (GPS) (Varnum and Chaffee, 1982), then the tracking station coordinates are held fixed and the main unknowns are the satellite starting elements and polar motion values. For geodynamic and geophysical applications such as crustal movements, earth tides, and geopotential models, the solution can include tracking station coordinates, Love numbers and various spherical harmonic coefficients. Coordinate reference frame studies involve the determination of polar motion and earth rotation values.

The numerical integration must be performed in an inertial (non-rotating) reference frame (IF). However, the tracking station coordinates and a number of components of the force model are known in an earth-fixed (EF) frame. This chapter describes the different reference frames and the relationships between them (Section 2.2). The details of the constituents of the force model are also given (Section 2.3). The numerical integration of the equations of satellite motion is treated in Section 2.4, and the formulation and solution of the least squares observation equations for the various unknowns in the OD problem is described in Section 2.5.

2.2 COORDINATE REFERENCE FRAMES AND TIME SCALES

2.2.1 Reference Frames and Time

The International Astronomical Union (IAU) has recently (1976 and 1979) adopted a set of new resolutions regarding constants, time scales and the new fundamental astronomical reference frame FK5. These resolutions apply as from 1 January 1984. They include a new equation for Greenwich Mean Sidereal Time (GMST) in terms of Universal Time (UT1), and new precession and nutation models (Kaplan, 1981).

The standard epoch in FK5 is J2000.0 or January 1^d.5 of the year 2000, which corresponds to Julian ephemeris date 2451545.0. For the purposes of orbit determination the inertial frame (IF) is a right-handed coordinate system (see Appendix B.1) with its origin at the earth's mass centre and it is defined by the mean equator and equinox of J2000.0. This is the frame in which the numerical integration is performed. The fundamental ephemeris of solar system bodies, which has been recomputed following the introduction of the new resolutions and is now known as Development Ephemeris Number DE200/LE200 (Melbourne, 1983), is also in this reference frame, but with the origin at the heliocentre.

The tracking station coordinates, as well as the geopotential field, are given in terms of an earth-fixed (EF) reference frame. This is also right-handed and is defined with its origin at the geocentre, its Z-axis directed towards the CIO pole as maintained by the Bureau International de l'Heure (BIH), and the X-axis towards the BIH zero meridian. The complete procedure of transforming IF to EF coordinates is explained in Sections 2.2.2 and 2.2.3 and summarized in Section 2.2.4.

It is appropriate to define here the time scales that are used in the OD process. These are required in order to compute the various arguments needed for the coordinate transformations.

Greenwich Apparent Sidereal Time (GAST) is the hour angle between the Greenwich meridian and the true equinox of date. Local Apparent Sidereal Time (LAST), the local hour angle of the true equinox of date, is determined by observations of stars at a number of observatories and is related to GAST by

$$\text{LAST} = \text{GAST} + \Lambda \quad (2.1)$$

where Λ : astronomical longitude of local meridian, measured positive east of Greenwich, in units of time.

The true equinox of date is derived by correcting for the effects of precession and nutation (see Section 2.2.2). Greenwich Mean Sidereal Time (GMST) is the Greenwich hour angle of the mean equinox of date, which is derived by correcting for the effects of precession only.

Universal Time (UT) is the time scale that is most closely related to the diurnal motion of the sun. It is determined by the observations of stars at a number of observatories (which make observations of their LAST) and comprises four slightly different scales denoted by UTO, UT1, UT2 and UTC. UTO is calculated directly from the observed sidereal time and has periodic and irregular variations due to polar motion and the variable rate of rotation of the earth (see Section 2.2.3). After each individual UTO has been corrected for polar motion, a weighted mean of all participating observatories results in the time scale known as UT1, which is maintained by BIH.

This is the navigator's and astronomer's time scale, as it represents a global determination of the rotation of the earth. It is connected to GMST by a simple expression ((2.20)). UT2 is derived by correcting UT1 for predicted values of seasonal variations in the rotation rate of the earth.

The advent of highly stable crystal and atomic clocks has led to the redefinition of the second as the basic unit of time by the International System of Units (SI) (Blair, 1974). The present definition is given in terms of a resonance of the cesium atom and it supersedes the previous ones which were given in the first place, in terms of the mean solar day and, until 1967, in terms of the earth's orbit. This has led to the establishment of an atomic time scale, namely International Atomic Time (TAI), which is based on the SI second. TAI is maintained by the BIH and represents the weighted mean of readings of several atomic clocks functioning in various establishments. However, TAI does not correspond with UT and there is a requirement for a time scale which is based on the SI second but which also keeps pace with changes in the earth's rotation rate. This time scale is Coordinated Universal Time (UTC) and is such that it differs from TAI by an integer number of seconds, but is adjusted by whole second steps to keep within 0.9 seconds of UT1. These adjustments are generally not predictable. The current value of TAI-UTC, since July 1 1983, is 22 seconds. The BIH, in its monthly Circular D and in its Annual Report (BIH, 1980), publishes values of UT1-UTC and UT1-TAI as functions of Universal Time.

UTC is the time scale that is most generally available through television, radio and satellite transmissions (Blair, 1974). For this reason it is specially suited to be the primary time reference for an OD process. Most satellite tracking stations record their observations in terms of UTC. Also, except for the leap seconds, UTC time intervals are constant and can be used to define the step size of the numerical integration procedures (Section 2.4.2).

The IAU resolutions (Kaplan, 1981) define two new time scales. Terrestrial Dynamical Time (TDT) is defined as

$$\text{TDT} = \text{TAI} + 32^{\text{s}}.184 \quad (2.2)$$

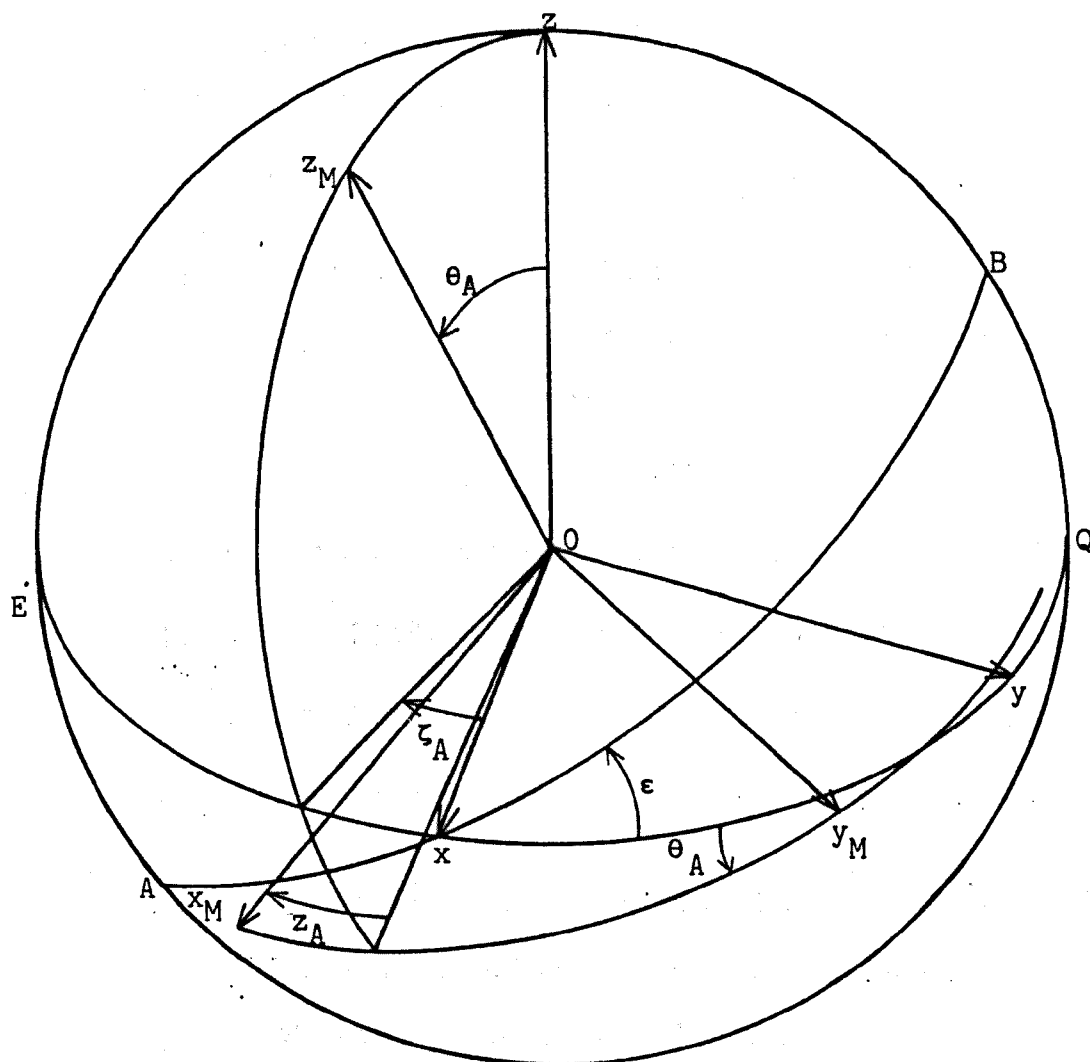
and is the time scale for an apparent geocentric ephemeris. It replaces Ephemeris Time (E.S.A.E, 1977), and continuity is maintained with the latter. Barycentric Dynamical Time (TDB) is the time scale for the equations of motion relative to the Solar System's barycentre. It differs from TDT by periodic relativistic terms. The new precession and nutation models are given in terms of TDB.

2.2.2 Precession and Nutation

Several of the concepts that are discussed in this section are illustrated in fig 2.II. This shows the celestial sphere at J2000.0. EQ is the trace of the mean equatorial plane on the celestial sphere. AB is the trace of the ecliptic, which is the plane of the earth's orbit. The intersection of the equatorial and ecliptic planes is the mean equinox of J2000.0 and defines the x - coordinate of the IF. The z - axis is defined by the pole of the equator through the geocentre, and the y - axis completes the right-handed system. So, the set of axes, x, y and z, represents the 'Mean of J2000.0' or IF.

The angle between equator and ecliptic is known as the obliquity of the ecliptic, ϵ .

The attraction of the moon, sun and planets on the earth causes the equator and the ecliptic, and hence the equinox, to be in a state of constant motion. The motion of the equator (or the celestial pole) is due to the attraction of the moon and sun on the earth's equatorial bulge. This consists of two components: the first is a westerly motion of the mean pole of the equator round the pole of the ecliptic, known as luni-solar precession. The second component, nutation, is a short-period, rather irregular motion of the true (actual) pole round the mean pole, with a main period of 18.6 years and an amplitude of about 9". The motion of the ecliptic is due to the action of the planets on the earth as a whole, and results in an eastward motion of the equinox of about 12" per century and a decrease of the obliquity of about 47" a century, known as planetary precession. The combined effects of luni-solar and planetary precession, known as general precession, are described by three angles, the equatorial precession parameters (Kaplan, 1981) ζ_A , z_A and θ_A . These are illustrated in fig.2.II and they connect the IF (Mean of J2000.0) to the Mean of Date reference frame. For the purposes of OD, this date would be the epoch of the integration step or that of the particular observation to the satellite. The transformation from IF to Mean of Date coordinates at a certain TDB time, t_{TDB} , is given by



EQ : equator at J2000.0

AB : ecliptic at J2000.0

ϵ E : obliquity of the ecliptic

x, y, z : IF (Mean of J2000.0) coordinate axes

x_M, y_M, z_M : Mean of Date coordinate axes

θ_A, ζ_A, z_A : equatorial precession parameters

Fig 2.II The Celestial Sphere and Precession

$$\underline{r}_M = Q \underline{r} \quad (2.3)$$

where $\underline{r} : (x, y, z)^T$, the IF coordinates

$\underline{r}_M : (x_M, y_M, z_M)^T$, the Mean of Date coordinates

and Q is the precession matrix :

$$Q = R_3(-z_A) R_2(\theta_A) R_1(-\zeta_A) \quad (2.4)$$

with $R_1(\varphi)$, $R_2(\varphi)$, $R_3(\varphi)$: rotation matrices for anticlockwise rotations through an angle φ about the x , y and z axes, respectively (see Appendix A)

The equatorial precession parameters are given by (Kaplan, 1981)

$$\begin{aligned} \zeta_A &= 2306''.2181 T + 0''.30188 T^2 + 0''.017998 T^3 \\ z_A &= 2306''.2181 T + 1''.09468 T^2 + 0''.018203 T^3 \\ \theta_A &= 2004''.3109 T - 0''.42665 T^2 - 0''.041833 T^3 \end{aligned} \quad (2.5)$$

where T : the interval, measured in Julian centuries of TDB, between J2000.0 and the required epoch (of the Mean of Date reference system), given by

$$T = (J - 2451545.0) / 36525 \quad (2.6)$$

and J is the TDB Julian date of the epoch.

Normally the UTC time, t_{UTC} , will be known and t_{TDB} is required. t_{UTC} is first converted to TAI, using the BIH values of UT1 - UTC and UT1 - TAI,

$$t_{TAI} = t_{UTC} + (UT1 - UTC) - (UT1 - TAI) \quad (2.7)$$

This, in turn, is converted to TDT (t_{TDT}) using (2.2). The TDB time is now given by

$$t_{\text{TDB}} = t_{\text{TDT}} + 0.001658 \sin(g + 0.0167 \sin g) \quad (2.8)$$

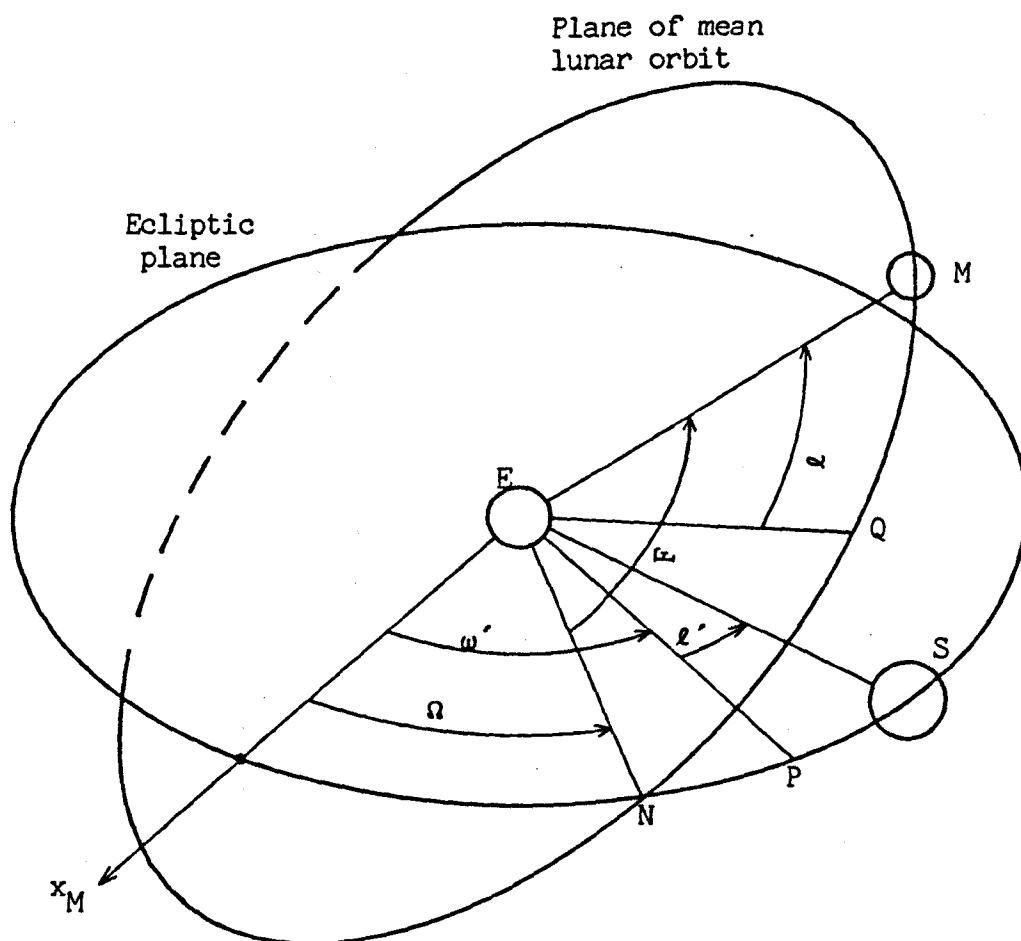
where g : mean anomaly of the earth in its orbit,

$$g = (357.528 + 35999.050 T) \times 2\pi/360^\circ \quad (2.9)$$

From t_{TDB} and the known date, the Julian date, J , is computed and, from this, the values of the precession parameters are evaluated. The interval, T , as defined in (2.6), is also used in the computation of the nutation arguments. The above expressions for the equatorial precession parameters have been adopted by the General Assembly of the IAU at its 1976 meeting. They represent the first change since 1896, when Newcomb's values were adopted (E.S.A.E., 1977).

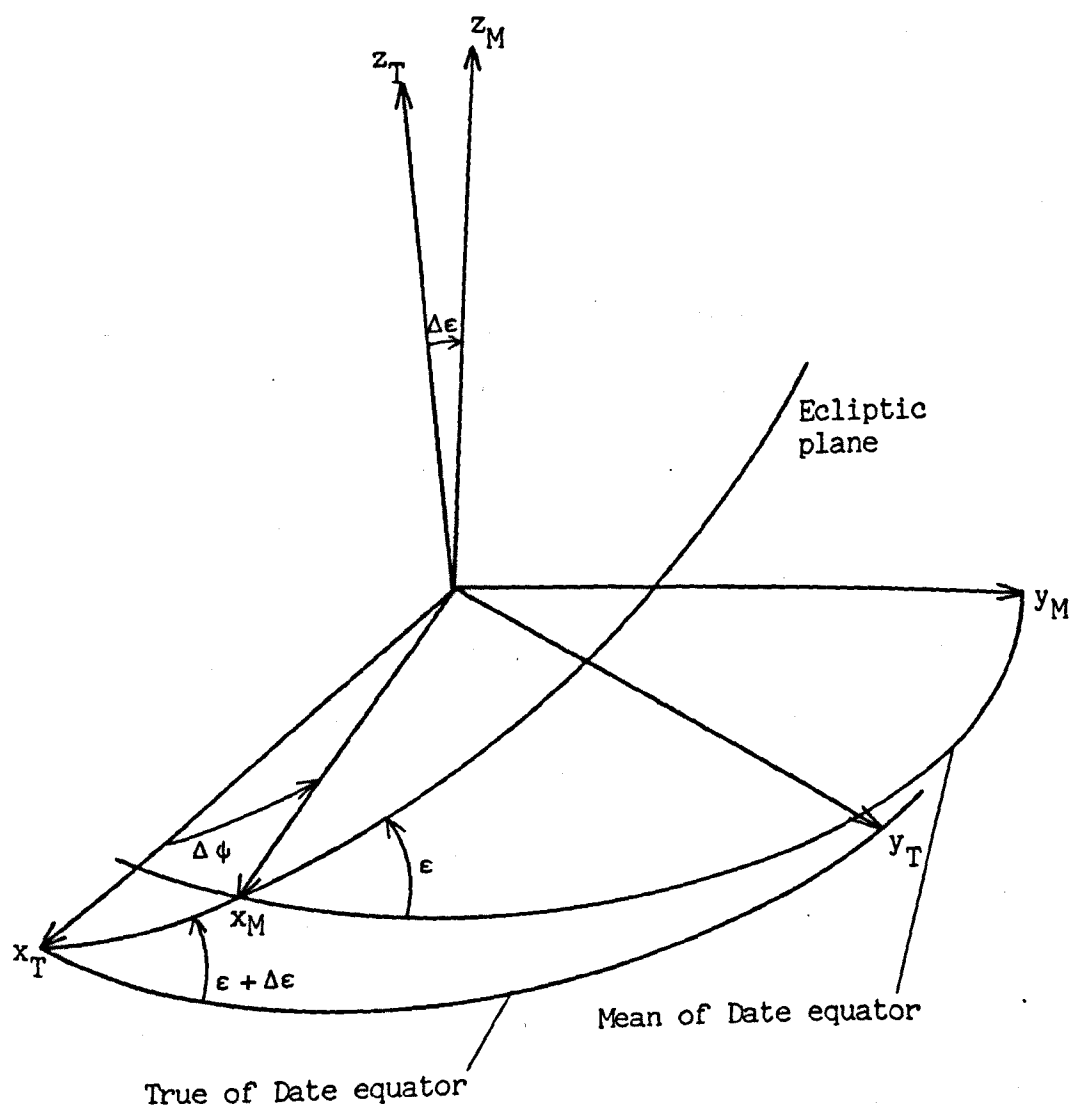
As mentioned above, nutation is the relatively short-period motion of the true pole about the mean pole. It varies with the position of the moon and the sun in their orbits round the earth. The principal term of nutation depends on the longitude of the ascending node of the moon's mean orbit on the ecliptic (Ω) measured from the mean equinox of date, and has a period of 18.6 years and an amplitude of $9''.2025$. Other terms arise, which depend on the mean anomalies of the moon and the sun (ℓ and ℓ'), on the moon's mean elongation from the node (F), and on the mean elongation of the moon from the sun (D). These fundamental arguments (ℓ , ℓ' , F , D and Ω) are illustrated in fig.2.III.

Nutation is described in terms of two angles, the nutation in longitude, $\Delta\psi$, and nutation in obliquity, $\Delta\epsilon$. These connect the Mean of Date system, as defined above, to the True of Date reference frame with coordinate axes x_T , y_T and z_T , as defined by the true equator and equinox of date. The relationship between the two systems is illustrated in fig.2.IV. The nutation matrix, N , is given by



- E : earth
- M : mean moon
- S : mean sun
- N : moon's ascending node
- L : moon's longitude ($= \Omega + F$)
- ω' : longitude of perihelion
- L' : sun's longitude ($= \omega' + l'$)
- D : moon's mean elongation from sun ($= L - L'$)
- P : perihelion
- Q : perigee of moon's orbit
- x_M : mean equinox of date

Fig 2.III The Fundamental Arguments



$\Delta\epsilon$: nutation in obliquity

$\Delta\phi$: nutation in longitude

x_T, y_T, z_T : True of Date coordinate axes

Fig 2.IV The Nutation Angles

$$N = R_1 (-\epsilon - \Delta\epsilon) R_3 (-\Delta\psi) R_1 (\epsilon) \quad (2.10)$$

and the transformation from Mean of Date to True of Date coordinates is given by

$$\underline{r}_T = N \underline{r}_M \quad (2.11)$$

where $\underline{r}_T : (x_T, y_T, z_T)^T$, the True of Date coordinates.

The nutation theory currently in use is the 1980 IAU Theory of Nutation (Kaplan, 1981). This replaces the theory developed by E. W. Woollard (E.S.A.E., 1977). The new theory was developed by J. Wahr (Wahr, 1981b) based on work done by H. Kinoshita and on geophysical model 1066A by F. Gilbert and A. Dziewonski (Melbourne, 1983). It uses a non-rigid model of the earth, without axial symmetry, which includes the effect of a solid inner core and a liquid outer core. Also, the reference pole is selected so that there are no diurnal motions of this pole with respect to an earth-fixed or a space-fixed reference frame. These are included implicitly in the new nutation theory, which takes into account all externally forced motions of the earth's spin axis. The nutations in longitude ($\Delta\psi$) and in obliquity ($\Delta\epsilon$) are given by a series of 106 terms as

$$\begin{aligned} \Delta\psi &= \sum_{i=1}^{106} (p_i + q_i T)'' \sin(a_i \ell + b_i \ell' + c_i F + d_i D + e_i \Omega) \\ \Delta\epsilon &= \sum_{i=1}^{106} (r_i + s_i T)'' \cos(a_i \ell + b_i \ell' + c_i F + d_i D + e_i \Omega) \end{aligned} \quad (2.12)$$

where T is as defined in (2.6) and

a_i, b_i, c_i, d_i, e_i : integer multiples of the fundamental arguments

$(p_i + q_i T)''$: coefficient of sine argument (in seconds of arc)

$(r_i + s_i T)''$: coefficient of cosine argument (in seconds of arc).

All the terms (a_i , b_i , etc) needed to compute $\Delta\phi$ and $\Delta\epsilon$ are tabulated in Kaplan (1981) and in Melbourne (1983).

The fundamental arguments are given by:

$$\ell = 485866.^{\circ}733 + (1325^{\circ} + 715922.^{\circ}633)T + 31.^{\circ}310 T^2 + 0.^{\circ}064 T^3 \quad (2.13)$$

$$\ell' = 1287099.^{\circ}804 + (99^{\circ} + 1292581.^{\circ}224)T - 0.^{\circ}577 T^2 - 0.^{\circ}012 T^3 \quad (2.14)$$

$$F = 335778.^{\circ}877 + (1342^{\circ} + 295263.^{\circ}137)T - 13.^{\circ}257 T^2 + 0.^{\circ}011 T^3 \quad (2.15)$$

$$D = 1072261.^{\circ}307 + (1236^{\circ} + 1105601.^{\circ}328)T - 6.^{\circ}891 T^2 + 0.^{\circ}019 T^3 \quad (2.16)$$

$$\Omega = 450160.^{\circ}280 - (5^{\circ} + 482890.^{\circ}539)T + 7.^{\circ}455 T^2 + 0.^{\circ}008 T^3 \quad (2.17)$$

where $1^{\circ} = 360^{\circ}$.

The obliquity of the ecliptic is

$$\epsilon = 84381.^{\circ}448 - 46.^{\circ}8150 T - 0.^{\circ}00059 T^2 + 0.^{\circ}001813 T^3 \quad (2.18)$$

2.2.3 Earth Rotation and Polar Motion

As mentioned earlier (Section 2.2.1), GMST, the Greenwich hour angle of the mean equinox of date, can be expressed in terms of UT1. This is because they are both measures of the earth's rotation, relative to the mean sun in the case of UT1, and to the fixed stars in the case of GMST. The UT1 time, t_{UT1} , is obtained from the known t_{UTC} , using the BIH value of UT1-UTC:

$$t_{UT1} = t_{UTC} + (UT1 - UTC) \quad (2.19)$$

GMST at 0^h UT1 of the day in question is given by

$$\begin{aligned} \text{GMST}(0^h \text{ UT1}) &= 24110.^s54841 + 8640184.^s812866 T_U \\ &\quad + 0.^s093104 T_U^2 - 6.^s2 \times 10^{-6} T_U^3 \end{aligned} \quad (2.20)$$

where T_U : number of centuries of 36525 days of UT
 elapsed since 2000 January 1, 12^h UT1
 (JD 2451545.0 UT1).

The exact GMST at the required epoch is

$$\text{GMST} = \text{GMST}(0^h \text{UT1}) + \Delta \text{GMST} \quad (2.21)$$

where

$$\begin{aligned} \Delta \text{GMST} = & (1.002737909350795 + 5.9006 \times 10^{-n} T_U \\ & - 5.9 \times 10^{-15} T_U^2) t_{\text{UT1}} \end{aligned} \quad (2.22)$$

and t_{UT1} : UT1 time elapsed from 0^h to exact UT1 epoch.
 (i.e. UT1 time of day)

GAST, the Greenwich hour angle of the true equinox of date, is
 computed from GMST by

$$\text{GAST} = \text{GMST} + \Delta\phi \cos \epsilon \quad (2.23)$$

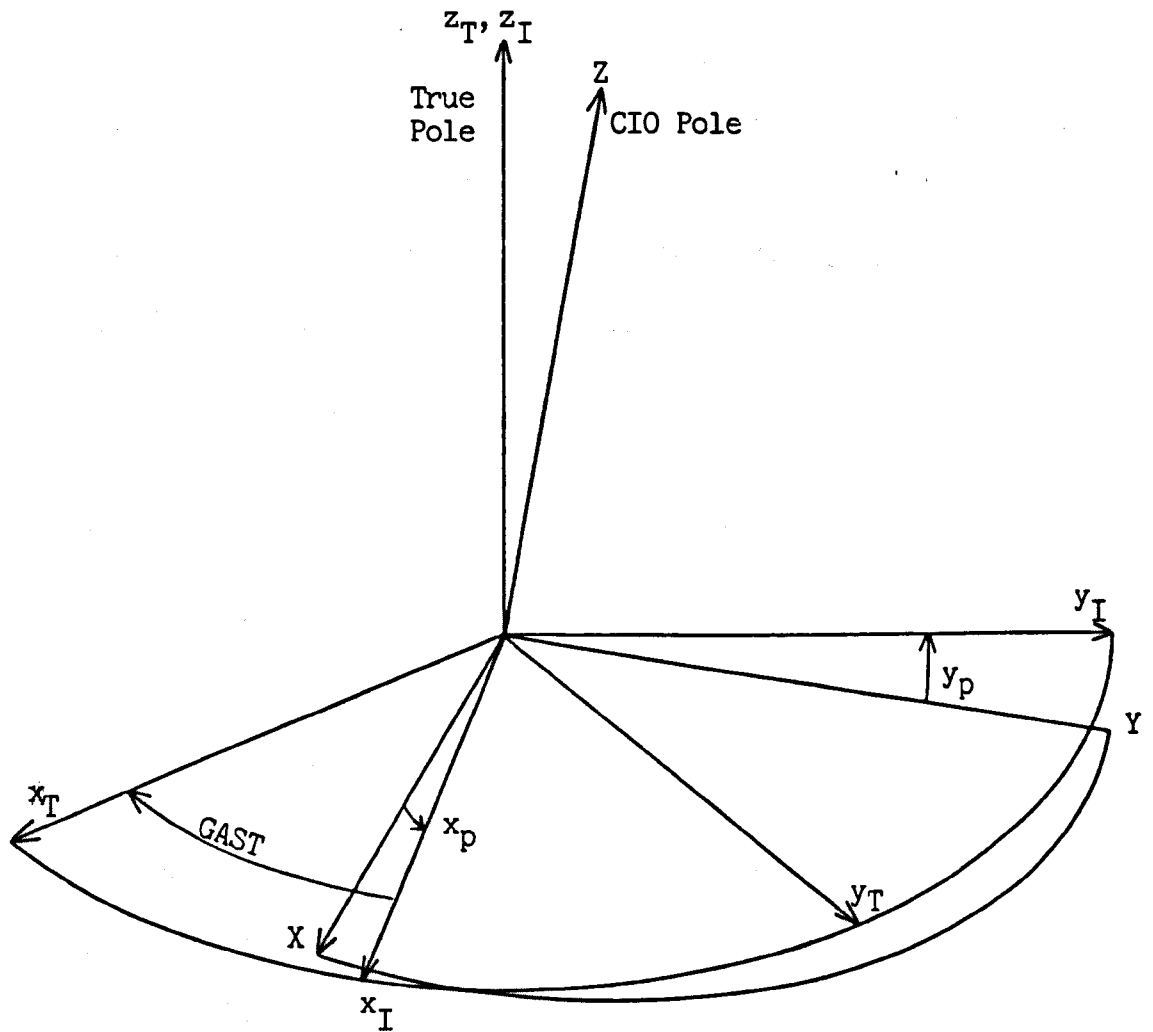
The transformation from a space-fixed reference frame to an earth-fixed reference frame is by means of sidereal time. As shown in fig.2.V, an anti-clockwise rotation about the z_T -axis (true pole) through GAST transforms the True of Date coordinates, x_T, y_T, z_T , to the instantaneous terrestrial (IT) reference frame. The earth rotation matrix, E , is defined as

$$E = R_s(\text{GAST}) \quad (2.24)$$

and the transformation to IT is given by

$$r_I = E r_T \quad (2.25)$$

where r_I : $(x_I, y_I, z_I)^T$, the IT coordinates.



x_p, y_p : polar motion angles
 x_I, y_I, z_I : Instantaneous Terrestrial coordinate axes
 x, y, z : EF coordinate axes

Fig 2.V Earth Rotation and Polar Motion

The IT reference frame is not truly earth-fixed. The true pole (instantaneous spin axis) moves relative to the body of the solid earth, an effect called polar motion. It is caused by a number of factors, the main ones being the non-parallelism of the earth's axis of rotation with that of maximum inertia (Chandler wobble) and meteorological effects (Bomford, 1980). This movement of the true pole is described by means of two angles, x_p and y_p , relative to a mean axis known as Conventional International Origin (CIO), as shown in fig.2.V. The polar motion matrix is

$$P = R_2(-x_p) R_1(-y_p) \quad (2.26)$$

and the transformation from IT to EF is given by

$$\underline{R} = P \underline{r}_I \quad (2.27)$$

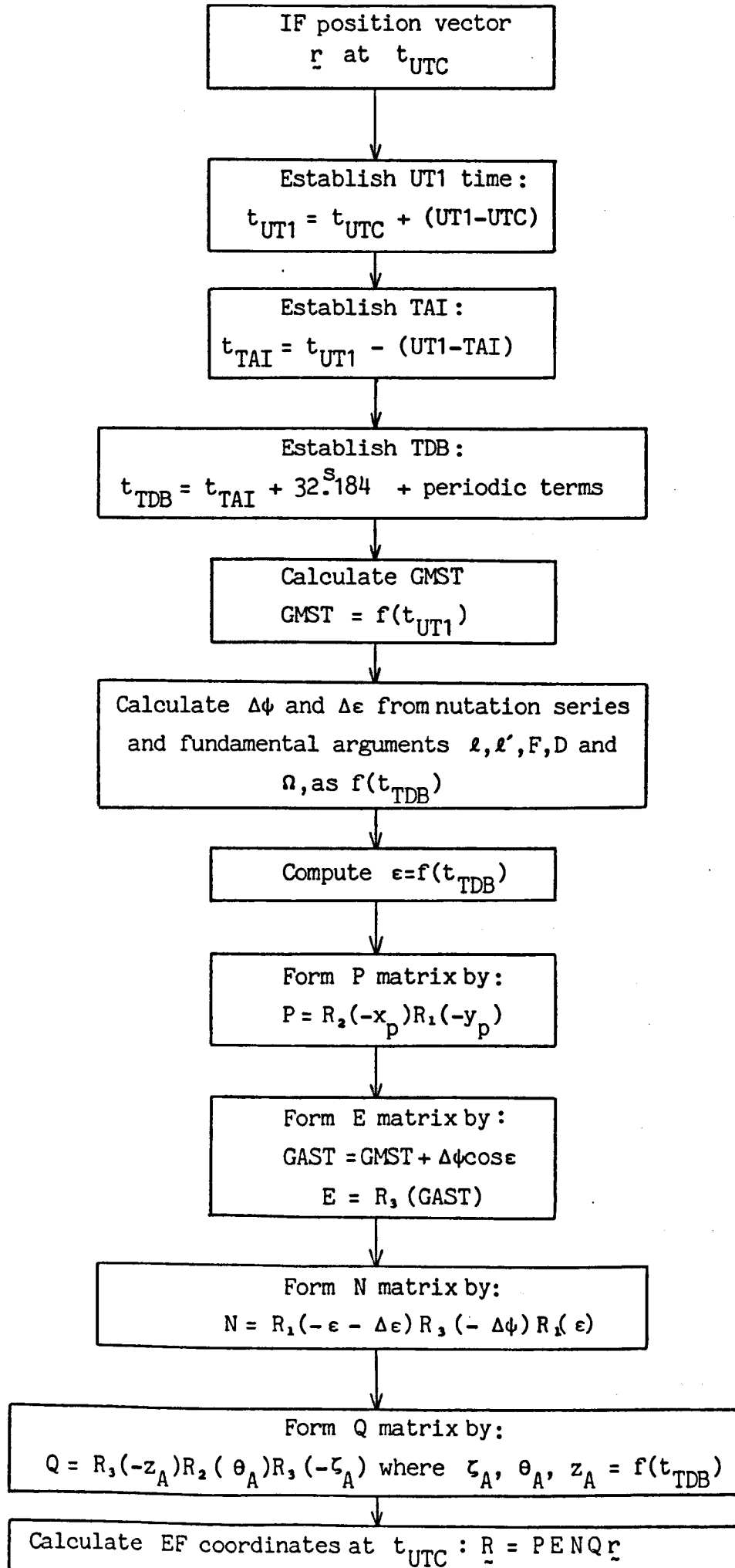
where $\underline{R} : (X, Y, Z)^T$, the earth-fixed coordinates.

This EF frame is the system in which terrestrial coordinates are given. CIO is defined as the mean spin axis of the years 1900 - 1905. It was originally established by the five observatories of the International Latitude Service (ILS) Mizasawa (Japan), Kitab (USSR), Carloforte (Italy), Gaithersburg and Ukiah (USA). The adopted latitudes of these observatories, located on the $38^{\circ}8'N$ parallel, define CIO. The polar motion values are obtained using techniques of optical astrometry (Bomford, 1980) by the International Polar Motion Service (IPMS), which publishes values of x_p and y_p based on observations of the ILS stations. It also receives data and publishes values from a number of other (about 50) observatories. The BIH, in its role of determining Universal Time, also derives values of polar motion. Its origin of coordinates was made to coincide with CIO

in 1968. Since 1973, the BIH values of x_p and y_p include a contribution due to Doppler observations of Transit satellites (McCarthy, 1982). The Doppler pole coordinates are derived through an OD process, in a solution which also determines the satellite starting elements (see Section 2.5), by the United States Defence Mapping Agency (DMA). The Doppler observations along with observations from other new techniques, such as Lunar Laser Ranging (LLR), have suggested that systematic errors existed in the 1968 BIH system. Subsequently the 1979 BIH system was introduced, which used Doppler observations taken between 1972 and 1978 to eliminate systematic effects of the classical astrometric methods. In 1980 the contribution of Transit Doppler tracking was 30% of the BIH polar motion determination. The values of polar motion and earth rotation (UT1 - UTC) are published monthly, with a delay of two months, in the BIH Circular D, and yearly, in the Annual Report. A weekly BIH Rapid Service is also maintained. The Rapid Service values are reproduced in the U.S. Naval Observatory (USNO) Time Service Announcement, Series 7, along with those of the USNO.

With the advent of new precise techniques such as Satellite Laser Ranging (SLR), Lunar Laser Ranging (LLR) and Very Long Baseline Interferometry (VLBI), it has been found necessary to review the methods currently in use in the determination of the Earth Rotation Parameters (ERP's). Project MERIT is a program of international collaboration to Monitor Earth Rotation and to Intercompare Techniques of observation and analysis (Wilkins, 1980). It is based on a proposal prepared by an IAU working group, that was endorsed by the International Union of Geodesy and Geophysics (IUGG), and its aim is to improve our knowledge and understanding of the variations in earth

rotation and polar motion. A short, preliminary observational campaign was held between August and October 1980 (Wilkins and Feissel, 1982). The main MERIT campaign is being held between September 1, 1983 and October 31, 1984, with an intensive phase between April 1 and June 30. A Joint Working Group of the International Association of Geodesy (IAG) and the International Astronomical Union (IAU) has also been set up, on the Establishment and Maintenance of a Conventional Terrestrial Reference System (COTES) (Mueller, 1983). The aims of this group are to prepare a proposal for the establishment of a new conventional terrestrial reference system and for the necessary international service(s) needed to maintain it. It was agreed to consider ways in which Project MERIT could be utilized to assist in the establishment of this system. (Wilkins and Feissel, 1982). Specifically, during the MERIT main campaign collocation of the VLBI, SLR and LLR systems will be attempted, with the help of mobile systems wherever possible, in order to detect systematic differences between the various IF and EF references frames which are inherent in these systems. Also, especially during the intensive phase of MERIT, it will be attempted to identify whether the various systems can detect short-period variations in the ERP's (x_p , y_p and $UT1 - UTC$) and, if so, how accurately. Until new services are adopted, the BIH is to serve as an interim system for all observational techniques.

Fig.2.VI Coordinate References Frames and Time Scales

2.2.4 Coordinate Transformations

The flow chart in fig.2.VI summarizes the procedure of transforming an IF position vector, \underline{r} , to the corresponding EF coordinates, \underline{R} , at a UTC time, t_{UTC} . The complete transformation, through a series of successive rotations, is

$$\underline{R} = \underline{P} \underline{E} \underline{N} \underline{Q} \underline{r} \quad (2.28)$$

The various expressions required to evaluate the polar motion (P), earth rotation (E), nutation (N) and precession (Q) matrices, are given in Sections 2.2.2 and 2.2.3. The reverse transformation, from EF to IF, is

$$\underline{r} = \underline{Q}^T \underline{N}^T \underline{E}^T \underline{P}^T \underline{R} \quad (2.29)$$

2.3 FORCE MODEL COMPONENTS

2.3.1 Introduction

The orbit of an artificial satellite around the earth is governed by the forces acting on it. These are classified as gravitational, surface and propulsion forces. Gravitational forces comprise the attractions of the earth, moon, sun and planets, as well as effects due to solar and lunar solid earth tides and ocean tides. Surface forces act on the satellite's exterior and vary according to its cross-sectional area, shape, mass and nature of the surface. They include solar and albedo radiation, and air drag. Propulsion forces are caused by thrusters which can occasionally be activated in order to control the satellite's orbit. Additional empirical accelerations, which would take into account any unmodelled

forces, could be included as unknowns in the OD process.

A common example is an empirical along-track acceleration for satellites which are at too high an altitude to be affected by air drag.

All these forces are described in detail in the following paragraphs. The various models are the ones adopted for the MERIT Campaign (Melbourne, 1983). The sum of the individual components gives the resultant acceleration of the satellite. This must be computed in an inertial frame (IF), since the numerical integration must be performed in a reference system that is independent of the earth's rotation.

2.3.2 Earth Gravitational Attraction

The earth's gravity is the major component of the forces acting on the satellite. It is defined by a geopotential expansion, in terms of spherical harmonics, which gives the earth's external potential, U , as a function of EF coordinates:

$$U = \frac{GM}{R} \left[1 + \sum_{n=2}^{\infty} \sum_{m=0}^n \left(\frac{a}{R} \right)^n P_n^m(\sin \phi) (C_n^m \cos m\Lambda + S_n^m \sin m\Lambda) \right] \quad (2.30)$$

where G : universal gravitational constant

M : mass of the earth

a : earth's equatorial radius

R, Λ, ϕ : EF spherical polar coordinates (geocentric radius, longitude and latitude) of point where potential is computed (see Appendix B.2)

n, m : degree and order of spherical harmonic expansion

$P_n^m(\sin \phi)$: Legendre function (of $\sin \phi$)

C_n^m, S_n^m : spherical harmonic coefficients

The Legendre functions are known as zonal harmonics when $m = 0$, tesseral harmonics when $m < n$, and sectorial harmonics when $m = n$. They are given by (Bomford, 1980)

$$P_n^m = \frac{\cos^m \phi}{(2^n)n!} \cdot \frac{d^{(m+n)}}{d \sin \phi} [(\sin^2 \phi - 1)^n] \quad (2.31)$$

and can be computed by recurrence relationships, for the case of zonals as

$$P_n^0 = \frac{1}{n} [(2n-1) \sin \phi P_{n-1}^0 - (n-1) P_{n-2}^0] \quad (2.32)$$

with $P_0^0 = 1$ (2.33)

and $P_1^0 = \sin \phi$ (2.34)

for the case of tesserals as

$$P_n^m = P_{n-2}^m + (2n-1) \cos \phi P_{n-1}^{m-1} \quad (2.35)$$

with $P_1^1 = \cos \phi$

and for the case of sectorials as

$$P_n^m = (2n-1) \cos \phi P_{n-1}^{m-1} \quad (2.36)$$

There are currently various expansions of the geopotential which have been derived by using combinations of satellite observations and terrestrial gravity measurements. They generally consist of the spherical harmonic coefficients, C_n^m and S_n^m , usually given in their normalized forms, \bar{C}_n^m and \bar{S}_n^m . The relationship between a coefficient and its normalized form is

$$C_n^m = N_n^m \bar{C}_n^m \quad (2.37)$$

$$\text{and} \quad S_n^m = N_n^m \bar{S}_n^m \quad (2.38)$$

where N_n^m is the normalizing factor

$$N_n^m = \left[\frac{(n-m)! (2n+1) (2-\delta_{0m})}{(n+m)!} \right]^{\frac{1}{2}} \quad (2.39)$$

with δ_{0m} : the Kronecker delta defined by

$$\delta_{0m} = 1 \text{ for } m=0 \text{ and } \delta_{0m} = 0 \text{ for } m \neq 0 \quad (2.40)$$

The satellite EF acceleration vector, $\ddot{\mathbf{R}}$, is given by the gradient of the potential field at the satellite :

$$\ddot{\mathbf{R}} = \nabla U \quad (2.41)$$

The individual components of the acceleration vector (\ddot{X} , \ddot{Y} , \ddot{Z}) are computed as

$$\ddot{R}_i = \frac{\partial U}{\partial R_i} = \frac{\partial U}{\partial R} \frac{\partial R}{\partial R_i} + \frac{\partial U}{\partial \Lambda} \frac{\partial \Lambda}{\partial R_i} + \frac{\partial U}{\partial \Phi} \frac{\partial \Phi}{\partial R_i} \quad (2.42)$$

where R_i : any one of the EF components of the satellite position vector (X, Y, Z)

\ddot{R}_i : the corresponding EF acceleration component

The partial derivatives of (2.42) are derived in Appendix C.1.

The satellite acceleration due to the earth's gravitational attraction is obtained in the IF reference frame, using

$$\ddot{\mathbf{r}} = \mathbf{Q}^T \mathbf{N}^T \mathbf{E}^T \mathbf{P}^T \ddot{\mathbf{R}} \quad (2.43)$$

where $\ddot{\mathbf{r}}$: $(\ddot{x}, \ddot{y}, \ddot{z})^T$, the IF acceleration vector

and the matrices \mathbf{P} , \mathbf{E} , \mathbf{N} and \mathbf{Q} are derived in Section 2.2.

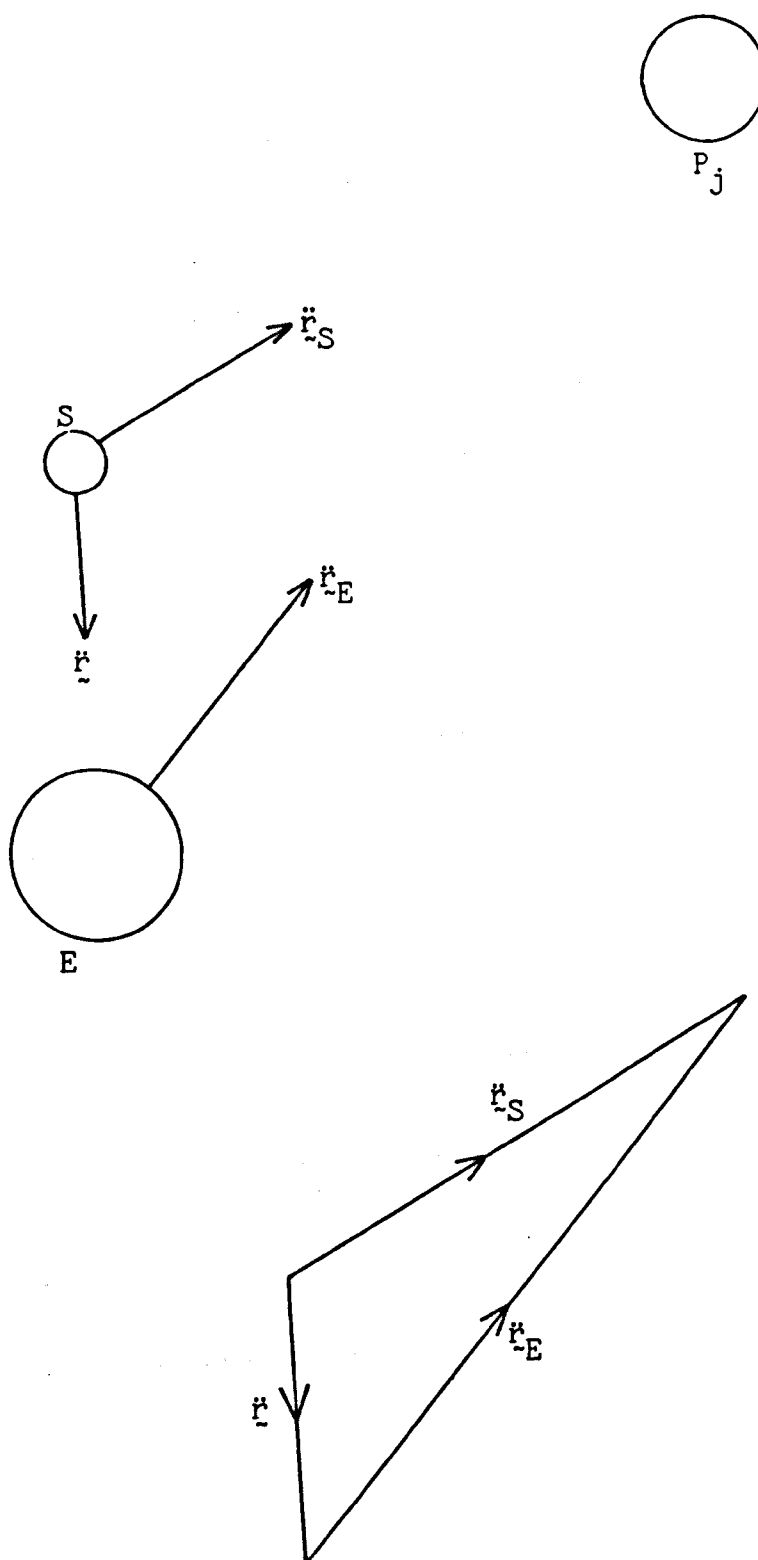
The best known geopotential models to date are the European GRIM models (Balmino et al, 1983), the Naval Surface Weapons Center NSWC10E used for the Transit system (Leroy, 1982), and the various NASA Goddard Earth Models (GEM). In particular, GEM9 is based on observations of 30 satellites, and GEM10 on combining the GEM9 data with surface gravimetry (Lerch et al, 1979). Several models have been specially 'tailored' to suit particular satellite systems. GEM-L2 (Lerch et al, 1983), for example, combines all the data for GEM9 with 2.5 years of SLR observations of the Laser Geodynamics Satellite (LAGEOS). The most recent NNSS model, NSWC10E-1, is the result of several revisions (Kumar, 1982) as more observations of the Transit satellites became available.

All the expansions of the geopotential truncate the infinite series of (2.30) after a finite number of terms. The OD process can determine these terms by introducing the spherical harmonic coefficients as unknowns in the least squares adjustment (see Section 2.5 and Appendix C.2). The earth's gravitational constant, GM , and equatorial radius, a , can also be determined.

2.3.3 Moon, Sun and Planetary Attractions

The moon, sun and planets of our solar system exert a gravitational pull on the satellite. This results in an acceleration component, $\ddot{\mathbf{r}}_S$, in a direction towards the 'third body', P_j , as shown in fig 2.VII. The earth, E , is also attracted by this third body (moon, sun or any planet), resulting in an acceleration, $\ddot{\mathbf{r}}_E$, of the earth towards P_j . The IF acceleration vector of the satellite (S) relative to the earth, caused by the third body attraction, is given by a 'triangle of forces' as (see fig 2.VII)

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_S - \ddot{\mathbf{r}}_E \quad (2.44)$$



Triangle of forces :

$$\vec{r} = \vec{r}_S - \vec{r}_E$$

Fig 2.VII Third Body Gravitational Attraction

The potential at the satellite due to P_j is

$$U_S = \frac{GM_j}{|\underline{r} - \underline{r}_j|} \quad (2.45)$$

where M_j : mass of third body

G : universal gravitational constant

$\underline{r}, \underline{r}_j$: IF position vectors of satellite and third body

This gives the acceleration of the satellite relative to P_j as

$$\ddot{\underline{r}}_S = \nabla U_S \quad (2.46)$$

resulting in

$$\ddot{\underline{r}}_S = \frac{-GM_j}{(|\underline{r} - \underline{r}_j|)^3} (\underline{r} - \underline{r}_j) \quad (2.47)$$

The potential at the earth arising from P_j is

$$U_E = \frac{GM_j}{|\underline{r}_E - \underline{r}_j|} \quad (2.48)$$

giving the acceleration of the earth relative to P_j

$$\ddot{\underline{r}}_E = \frac{-GM_j}{(|\underline{r}_E - \underline{r}_j|)^3} (\underline{r}_E - \underline{r}_j) \quad (2.49)$$

where \underline{r}_E : IF position vector of the earth

Since the earth's mass-centre is the origin of coordinates,

$$\underline{r}_E = 0 \quad (2.50)$$

so, (2.49) becomes

$$\ddot{\underline{r}}_E = \frac{GM_j}{r_j^3} \underline{r}_j \quad (2.51)$$

Substituting (2.47) and (2.51) in (2.44) gives

$$\ddot{\mathbf{r}} = -GM_j \left[\frac{\mathbf{r} - \mathbf{r}_j}{(|\mathbf{r} - \mathbf{r}_j|)^3} + \frac{\mathbf{r}_j}{r_j^3} \right] \quad (2.52)$$

where r_j : distance between third body and geocentre.

For the evaluation of the IF satellite acceleration from (2.52), the coordinates and masses of the bodies in the solar system are required. The positions of the moon and the planets are given in a heliocentric IF, for 0^h.0 TDB of each day, in Development Ephemeris Number DE200/LE200 (Kaplan, 1981). Associated with this ephemeris are a set of astronomical constants used in its creation. These include the masses of the planets relative to that of the sun, as well as geocentric and heliocentric gravitational constants. Firstly, all the coordinates must be converted to the geocentric IF by subtracting the coordinates of the earth from those of the other bodies. The position vector of the sun is now given as that of the earth, multiplied by -1. The coordinates of the required third body can then be interpolated, using a suitable interpolation scheme (see Appendix E), to the required epoch (of the numerical integration time step). Considerations of computational effort and required precision may result in a decision to disregard the effects of the more distant or less massive planets.

2.3.4 Solid Earth Tides

The potential at any point, A (see fig 2.VIII), on the surface of the earth (E) due to either the moon or the sun (P_j) is

$$U_A = \frac{GM_j}{\rho} \quad (2.53)$$

where M_j : mass of moon or sun.

The distance from P_j to A, ρ , can be computed using the cosine rule as

$$\rho^2 = r^2 + r_j^2 - 2r r_j \cos z \quad (2.54)$$

where r : distance from A to geocentre

r_j : distance from geocentre to centre of mass of moon or sun

The angle, z , is the angle AOP_j in fig 2.VIII and is approximately the zenith distance of the moon or sun at A. Equation (2.53) can be written as

$$U_A = \frac{GM_j}{(r^2 + r_j^2 - 2r r_j \cos z)^{\frac{1}{2}}} \quad (2.55)$$

This can be expressed as an infinite series in terms of Legendre functions, as

$$U_A = \frac{GM_j}{r_j} \sum_{n=0}^{\infty} \left(\frac{r}{r_j} \right)^n P_n(\cos z) \quad (2.56)$$

The tide raising potential is then

$$U_T = \frac{GM_j}{r_j} \sum_{n=2}^{\infty} \left(\frac{r}{r_j} \right)^n P_n(\cos z) \quad (2.57)$$

The only significant effect is that of the second harmonic ($n = 2$), giving

$$U_T = \frac{GM_j r^2}{r_j^3} P_2(\cos z) \quad (2.58)$$

The earth responds to this tidal potential by deforming, as shown by the dotted line in fig 2.VIII. It bulges out in directions towards and away from the moon (or sun), and caves in at right angles to these directions. The additional potential, due to the deformed shape, is

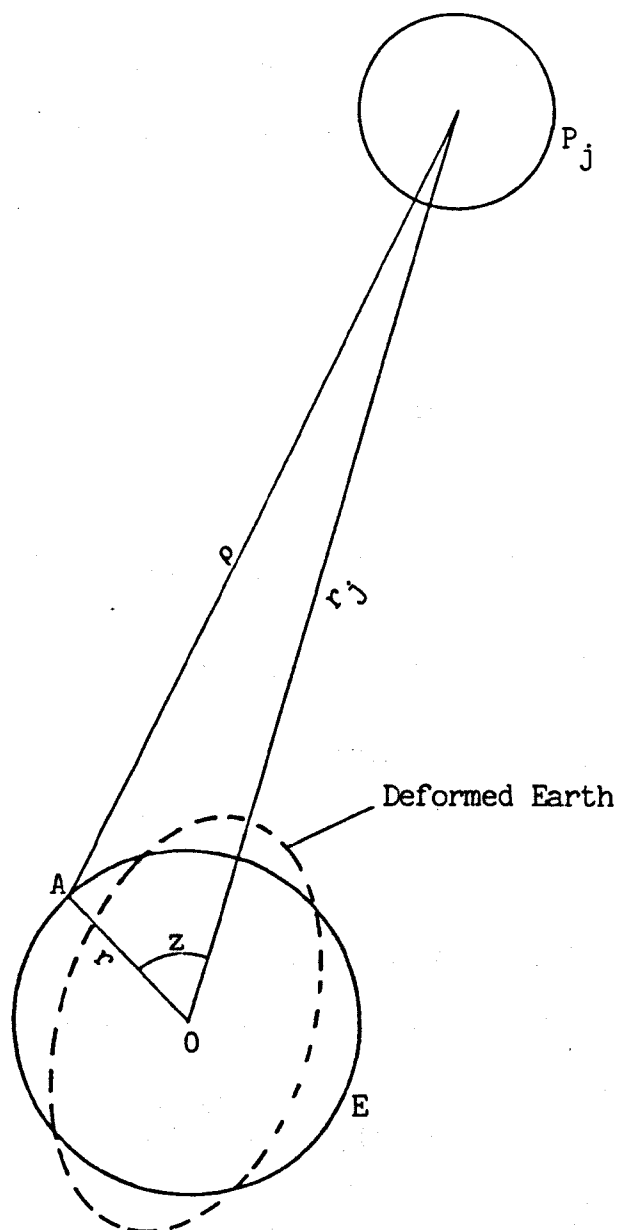


Fig 2.VIII Earth Tides

itself a second harmonic of the form

$$U = \frac{H}{r} \left(\frac{a}{r} \right)^2 P_2(\cos z) \quad (2.59)$$

where H : constant

a : earth's equatorial radius

r : distance from geocentre to point where potential is evaluated (e.g. at the satellite)

On the earth's surface $r = a$, giving

$$U_B = \frac{H}{a} P_2(\cos z) \quad (2.60)$$

The potential due to the deformed earth on the earth's surface is also given from (2.58) as

$$U_B = k_2 \frac{GM_j a^2}{r_j^3} P_2(\cos z) \quad (2.61)$$

where k_2 : Love number (nominal value = 0.3)

Equating (2.60) to (2.61) gives

$$H = k_2 \frac{GM_j a^3}{r_j^3} \quad (2.62)$$

and substituting this in (2.59) gives the potential of the tidal bulge at the satellite (or any other point), as

$$U = k_2 \frac{GM_j a^5}{r^3 r_j^3} P_2(\cos z) \quad (2.63)$$

The Legendre function, $P_2(\cos z)$, is given by

$$P_2(\cos z) = \frac{3}{2} \cos^2 z - \frac{1}{2} \quad (2.64)$$

with

$$\cos z = \frac{\mathbf{r} \cdot \mathbf{r}_j}{r r_j} \quad (2.65)$$

where \mathbf{r} : (x, y, z) , IF position vector of satellite

\mathbf{r}_j : (x_j, y_j, z_j) , IF position vector of moon (or sun)

The potential of the tidal bulge at the satellite becomes

$$U = k_2 \frac{GM_j a^5}{2r^3 r_j^3} \left[3 \frac{(\underline{r} \cdot \underline{r}_j)^2}{r^2 r_j^2} - 1 \right] \quad (2.66)$$

and the IF acceleration of the satellite is $\text{grad } U$ (∇U), giving

$$\ddot{\underline{r}} = k_2 \frac{GM_j a^5}{2r^3 r_j^3} \left[-15 \frac{(\underline{r} \cdot \underline{r}_j)^2}{r^2 r_j^2} \underline{r} + 6 \frac{(\underline{r} \cdot \underline{r}_j)}{r r_j} \underline{r}_j + 3 \underline{r} \right] \quad (2.67)$$

This acceleration is evaluated for both the moon and the sun, the coordinates of which are derived from DE200/LE200 as explained in Section 2.3.3.

The Love number, k_2 , is a measure of the response of the earth to the tidal potential, U_T . However, k_2 is not constant but varies for tides of different frequencies. The effect of the frequency dependent Love numbers is best included as corrections to the normalized spherical harmonic coefficients (\bar{C}_n^m and \bar{S}_n^m in (2.37) and (2.38)) of the geopotential expansion. A two-step procedure is therefore used to compute the effect of earth tides on the satellite. In the first step a nominal value of $k_2 = 0.3$ is used in equation (2.67). The second step involves the evaluation of corrections $\Delta \bar{C}_n^m$ and $\Delta \bar{S}_n^m$, due to the difference of the actual value of k_2 from the nominal one, for tides of different frequencies. The corrections for step 2 recommended for the MERIT campaign are (Melbourne, 1983)

$$\Delta \bar{C}_n^m = A_M \Sigma_{s(n,m)} \delta k_s H_s \begin{pmatrix} \cos \theta_s \\ \sin \theta_s \end{pmatrix} \begin{matrix} n+m \text{ even} \\ n+m \text{ odd} \end{matrix} \quad (2.68)$$

$$\Delta \bar{S}_n^m = A_M \Sigma_{s(n,m)} \delta k_s H_s \begin{pmatrix} -\sin \theta_s \\ \cos \theta_s \end{pmatrix} \begin{matrix} n+m \text{ even} \\ n+m \text{ odd} \end{matrix} \quad (2.69)$$

where δk_s : difference between Wahr model (Wahr, 1979 and 1981a) value of k_s and nominal value of $k_2 = 0.3$, in the sense $k_s - k_2$

H_s : amplitude (in metres) of tide s , taken from
Cartwright and Tayler (1971)

$$A_M = \frac{(-1)^m}{a(4\pi(2 - \delta_{om}))^{\frac{1}{2}}} \quad (2.70)$$

where δ_{om} : Kronecker delta

a : earth's equatorial radius

and $\theta_s = \underline{n} \cdot \underline{\beta} = \sum_{i=1}^6 n_i \beta_i \quad (2.71)$

where $\underline{\beta}$: (τ, s, h, p, N', p_1) , the vector of Doodson variables

\underline{n} : $(n_1, n_2, n_3, n_4, n_5, n_6)$ the vector of the integer multiples of the Doodson variables (Cartwright and Tayler, 1972)

The vector \underline{n} is sometimes given in coded form, known as Doodson's notation, as

$$D_n = d_1 d_2 d_3 . d_4 d_5 d_6 \quad (2.72)$$

where $d_1 = n_1$

and $d_i = n_i + 5$ for $i = 2, 3, 4, 5, 6$

The components of $\underline{\beta}$ are related to the fundamental arguments of nutation (see Section 2.2.2 and fig 2.III) by

$$s = F + \Omega$$

$$h = s - D$$

$$p = s - \ell \quad (2.73)$$

$$N' = -\Omega$$

$$p_1 = s - D - \ell'$$

$$\tau = \text{GMST} + \pi - s$$

Using a cut-off amplitude of 9×10^{-12} for the product $A_m \delta k_s H_s$ in (2.68) and (2.69) for tides of each frequency, only six diurnal tides (for $\Delta \bar{C}_2^1$ and $\Delta \bar{S}_2^1$) and two semi-diurnal tides (giving $\Delta \bar{C}_2^2$ and $\Delta \bar{S}_2^2$) need to be considered.

Apart from their effect on the satellite, solid earth tides also change the coordinates of the tracking station as the earth deforms. This movement can be up to 32 cm for lunar tides and 15 cm for solar tides, and has to be taken into account in precise orbit determination. The components of the IF displacement vector (d_p) of station p, in spherical coordinates, are

$$d_{rp} = \frac{h_2 U_T}{g} \quad (2.74)$$

$$d_{\lambda p} = \frac{l_2}{g \cos \varphi_p} \frac{\partial U_T}{\partial \lambda_p} \quad (2.75)$$

$$d_{\varphi p} = \frac{l_2}{g} \frac{\partial U_T}{\partial \varphi_p} \quad (2.76)$$

where g : acceleration due to gravity

$d_{rp}, d_{\lambda p}, d_{\varphi p}$: components of the IF displacement vector, d_p , of station p, in spherical coordinates

h_2 : Love number

l_2 : Shida number

$r_p, \lambda_p, \varphi_p$: IF position of station p expressed in spherical coordinates

From (2.58) it can be seen that

$$\frac{\partial U_T}{\partial r} = \frac{2 U_T}{r} \quad (2.77)$$

Substituting this in (2.74) gives

$$d_{rp} = \frac{h_2 r_p}{2g} \frac{\partial U_T}{\partial r_p} \quad (2.78)$$

The IF station displacement vector becomes

$$\begin{aligned} \underline{d}_p = & \frac{h_2 r_p}{2g} \left(\frac{\partial U_T}{\partial r_p} \hat{r}_p \right) + \frac{\ell_2 r_p}{g} \left(\frac{1}{r_p} \frac{\partial U_T}{\partial \varphi_p} \hat{\varphi}_p \right) + \frac{\ell_2 r_p}{g} \left(\frac{1}{r_p \cos \varphi_p} \frac{\partial U_T}{\partial \lambda_p} \hat{\lambda}_p \right) \\ & \dots \end{aligned} \quad (2.79)$$

where $\hat{r}_p, \hat{\lambda}_p, \hat{\varphi}_p$: unit vectors in the directions of r_p, λ_p and φ_p
Also, in spherical polar coordinates, the gradient of the tidal potential at p is

$$\begin{aligned} \nabla U_T = & \frac{\partial U_T}{\partial r_p} \hat{r}_p + \frac{1}{r_p} \frac{\partial U_T}{\partial \varphi_p} \hat{\varphi}_p + \frac{1}{r_p \cos \varphi_p} \frac{\partial U_T}{\partial \lambda_p} \hat{\lambda}_p \\ & \dots \end{aligned} \quad (2.80)$$

Combining (2.79) and (2.80), and substituting (2.77), gives

$$\underline{d}_p = \frac{\ell_2 r_p}{g} \nabla U_T + \frac{h_2 - 2\ell_2}{g} U_T \hat{r}_p \dots \quad (2.81)$$

Also,

$$g = \frac{GM}{a^2} \dots \quad (2.82)$$

The IF displacement of station p due to lunar or solar earth tides is now given by

$$\begin{aligned} \underline{d}_p = & \frac{M_j a^4}{M r_j^3} \left[3\ell_2 \frac{(\underline{r}_p \cdot \underline{r}_j)}{r_p r_j^2} \underline{r}_j + \left[3\left(\frac{h_2}{2} - \ell_2\right) \frac{(\underline{r}_p \cdot \underline{r}_j)^2}{(r_p r_j)^2} - \frac{h_2}{2} \right] \frac{\underline{r}_p}{r_p} \right] \\ & \dots \end{aligned} \quad (2.83)$$

The Love and Shida numbers (h_2 and l_2) vary for tides of different frequencies. If nominal values of 0.6090 and 0.0852 are used respectively, and a cut-off amplitude of 0.005m in radial displacement (the biggest effect) is required, the EF geocentric radius of the station has to be corrected by

$$\Delta R_p = -0.0253 \sin \phi_p \cos \phi_p \sin(\text{GMST} + \Lambda_p) \quad (2.84)$$

where R_p , Λ_p , ϕ_p : EF geocentric radius, longitude and latitude of p

ΔR_p : correction to geocentric radius, in metres

The above expression is for the K1 frequency, where h_2 from Wahr's theory (Wahr, 1979) is 0.5203, and this correction has a maximum amplitude of 0.013m.

Earth tides also affect the earth's rotation rate. This is because the tidal deformation changes its moments of inertia, thus changing UT1, the duration of the day and the earth's angular velocity, ω (Woolard, 1959). This effect has to be removed if earth rotation results obtained at different epochs are to be compared. The procedure for computing this correction is given in Melbourne (1983).

2.3.5 Ocean Tides

The surface of the open ocean is an equipotential surface, except for effects due to temperature, pressure, salinity and currents. This surface is in constant motion because of the tidal potential of equation (2.58). The ocean tide contains a large number of frequencies, and this fluctuating ocean level results in a variable load on the solid earth which responds by deforming. The additional potential outside the earth, arising from this deformed shape, can be expressed as a

series of spherical harmonics (Cazenave et al, 1977) as

$$U = 4\pi G a \rho_w \sum_{n=2}^{\infty} \sum_{m=0}^n \sum_s \frac{1 + k'_n}{2n+1} \left(\frac{a}{R}\right)^{n+1} (D_{s,n,m} \cos m\Lambda + E_{s,n,m} \sin m\Lambda) P_n^m$$

... (2.85)

where ρ_w : density of sea water

R, Λ, Φ : EF spherical coordinates of satellite (or any other point

k'_n : load deformation coefficients

a : earth's equatorial radius

$$D_{s,n,m} = (C_{s,n,m}^+ + C_{s,n,m}^-) \cos \theta_s + (S_{s,n,m}^+ + S_{s,n,m}^-) \sin \theta_s$$

... (2.86)

$$E_{s,n,m} = (S_{s,n,m}^+ - S_{s,n,m}^-) \cos \theta_s - (C_{s,n,m}^+ - C_{s,n,m}^-) \sin \theta_s$$

... (2.87)

where $C_{s,n,m}^{\pm}, S_{s,n,m}^{\pm}$: ocean tide coefficients for the tide
constituent, s

θ_s : argument of the tide, s , as defined in (2.71)

On comparing equation (2.85) with (2.30) for the geopotential, it can be seen that the ocean tide potential can be expressed as corrections to the normalized spherical harmonic coefficients, as

$$\Delta \bar{C}_n^m = \frac{4\pi a^2 \rho_w}{M N_n^m} \frac{1 + k'_n}{2n+1} \sum_s \Sigma_{s(n,m)} D_{s,n,m} \quad \dots (2.88)$$

$$\Delta \bar{S}_n^m = \frac{4\pi a^2 \rho_w}{M N_n^m} \frac{1 + k'_n}{2n+1} \sum_s \Sigma_{s(n,m)} E_{s,n,m} \quad \dots (2.89)$$

where M : earth's mass

N_n^m : normalizing factor as given in (2.39)

The $C_{s,n,m}^+$ and $S_{s,n,m}^+$ in (2.86) and (2.87) denote prograde waves, while the $C_{s,n,m}^-$ and $S_{s,n,m}^-$ denote retrograde waves. The effect of the retrograde waves can be ignored if the long period ($m=0$) prograde terms are doubled. The values of the coefficients, $C_{s,n,m}^+$ and $S_{s,n,m}^+$, needed to compute the ocean tide correction, are given in Melbourne (1983) for the (MERIT recommended) Schwiderski ocean tide model (Schwiderski, 1980). The adopted load deformation coefficients (k_n') to degree $n=6$ are also given, and corrections can be evaluated up to degree 6 and order $m=2$. Ocean tide coefficients are given for the long period tides ($m=0$) S_{sa} , M_m and M_f , for the diurnal tides ($m=1$) Q_1 , O_1 , P_1 and K_1 and for the semi-diurnal ($m=2$) N_2 , M_2 , S_2 and K_2 .

As for solid earth tides, there is an ocean tide effect on station coordinates. The magnitude of this on the height of the tracking stations is generally of the order of 1 cm, but it can have a range of up to 10 cm. Care must be taken to apply this correction if it is within the accuracy requirements of the OD process. It has been evaluated for 25 SLR and VLBI sites and is given in Melbourne (1983).

2.3.6 Empirical Accelerations

These accelerations are sometimes required in order to account for any deficiencies of the force model in specified directions. They are of the form

$$\ddot{\mathbf{r}} = C \hat{\mathbf{n}} \quad (2.90)$$

where C : constant

$\hat{\mathbf{n}}$: unit vector in required direction

The most common empirical accelerations are in the along-track,

$$\ddot{\underline{r}} = C_a \frac{\dot{\underline{r}}}{r} \quad (2.91)$$

in the radial

$$\ddot{\underline{r}} = C_r \frac{\underline{r}}{r} \quad (2.92)$$

and in the cross-track directions:

$$\ddot{\underline{r}} = C_c \frac{\underline{r} \times \dot{\underline{r}}}{|\underline{r} \times \dot{\underline{r}}|} \quad (2.93)$$

where \underline{r} , $\dot{\underline{r}}$, $\ddot{\underline{r}}$: IF satellite position, velocity and acceleration vectors

r , \dot{r} , \ddot{r} : magnitudes of these vectors

The coefficients C_a , C_r and C_c can be given known values, or they can be adjusted parameters in the least squares solution (see Section 2.5 and Appendix C.2).

2.3.7 Air Drag

The IF acceleration of the satellite arising from air drag is

$$\ddot{\underline{r}} = -\frac{1}{2} C_D \left(\frac{A}{m} \right) \rho_a \underline{v}_r \underline{v}_r \quad (2.94)$$

where C_D : satellite drag coefficient

A, m : satellite cross-sectional area and mass

ρ_a : air density at the satellite

\underline{v}_r : IF velocity vector of the satellite relative to the atmosphere

The velocity, \underline{v}_r , is given by a vector triangle as (see fig 2.IX)

$$\underline{v}_r = \dot{\underline{r}} - \underline{v}_a \quad (2.95)$$

where $\dot{\underline{r}} : (\dot{x}, \dot{y}, \dot{z})$ IF satellite velocity vector

\underline{v}_a : IF velocity vector of the atmosphere

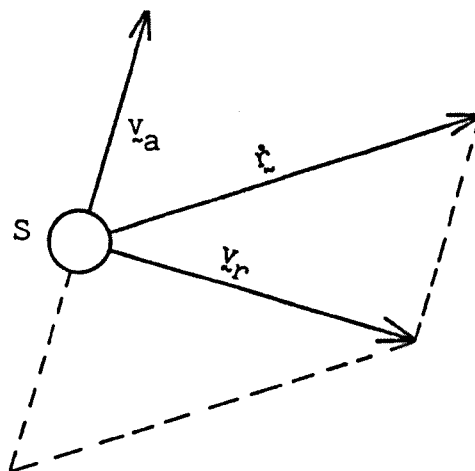


Fig 2.IX Satellite Velocity Relative to Atmosphere

The atmosphere is assumed to rotate at the same rate as the earth. This gives its True of Date (see Section 2.2.2) velocity at the satellite, as the vector cross product

$$\underline{v}_{Ta} = \underline{\omega} \times \underline{r}_T \quad (2.96)$$

where $\underline{\omega} : (0, 0, \omega)^T$ the earth's True of Date rotation rate vector

$\underline{r}_T : (x_T, y_T, z_T)^T$, the satellite True of Date coordinates.

This results in

$$\underline{v}_{Ta} = (-\omega y_T, \omega x_T, 0)^T \quad (2.97)$$

and the atmosphere's IF velocity vector, \underline{v}_a , needed in equation (2.95) is obtained by rotating the True of Date vector, using the nutation (N) and precession (Q) matrices, as

$$\underline{v}_a = Q^T N^T \underline{v}_{Ta} \quad (2.98)$$

A model is now required to give the air density (ρ_a) at the satellite. Such a model can be found in CIRA (1972) based mostly on work by L. G. Jachia. The atmospheric density is computed from the known date, satellite and sun positions, and solar flux and geomagnetic data. The model assumes that the atmosphere is composed entirely of nitrogen, oxygen (molecular and atomic), argon, helium and hydrogen. The density is given primarily as a function of height and temperature, and it is corrected for effects arising from geomagnetic activity and seasonal-latitudinal and semi-annual variations. The temperature is itself a function of height, solar and geomagnetic activity and diurnal variations.

CIRA (1972) gives a complete atmospheric density model for use at altitudes between 110 and 2000 km. It also includes a FORTRAN subroutine to evaluate this density. For altitudes higher than 2000 km there are currently no valid models, as the effect of air drag is very small. It can be best taken into account as an empirical along-track acceleration (see Section 2.3.6) with a scaling constant, C_a , that can be adjusted in the OD process. The drag coefficient, C_D , in equation (2.94) can also be a solve-for parameter when this equation is used (at altitudes less than 2000 km). The magnitude of the drag acceleration is directly proportional to the satellite's cross-sectional area-to-mass ratio (A/m). So the drag effect can be minimized if this ratio is kept small. On the other hand, for such cases as studies of the upper atmosphere it may be desirable to have a satellite that is significantly affected by air drag, in which case this ratio is maximized (e.g. balloon-type satellites).

Drag is classified as a surface force because it depends on the particular satellite, unlike gravitational forces which only depend on its position vector and time. The nature of the satellite's surface will determine C_D and, as mentioned above, the A/m ratio will determine its sensitivity to drag and other surface forces (see also Section 2.3.8). This ratio is a constant only for a spherical satellite or one whose orientation relative to the atmospheric velocity vector, v_a , is constant. So an added complication arises when the satellite has an awkward shape, especially when there are moving solar panels which have to be kept pointing towards the sun. Because of this, and of uncertainties in the existing air density models and in the value of C_D , drag, in common with other surface forces, cannot be modelled as accurately as the gravitational forces described earlier in this chapter.

2.3.8 Solar and Albedo Radiation

The sun constantly emits radiation whose intensity obeys the inverse square law with varying distance from its centre. At one astronomical unit (the distance equivalent to the semi-major axis of the earth's orbit) this intensity (power per unit area) is approximately

$$I_0 = 1367.2 \text{ W/m}^2 \quad (2.99)$$

and at the satellite, it is

$$I = I_0 \left[\frac{A}{\|\vec{r} - \vec{r}_j\|} \right]^2 \quad (2.100)$$

where A : one astronomical unit ($1.4959787 \times 10^{11} \text{ m}$)

\vec{r} : satellite IF position vector

\vec{r}_j : IF position vector of sun (from DE200/LE200)

Since force is power divided by velocity, the force per unit area (or solar radiation pressure) on the satellite due to incident radiation is

$$P_s = \frac{I}{c} \quad (2.101)$$

where c : speed of light in vacuo (2.99792458×10^8 m/s).

Multiplication by the ratio A/m (see equation (2.94)) converts the solar radiation pressure (P_s) into an acceleration whose direction is away from the sun. The IF satellite acceleration vector arising from direct solar radiation is then

$$\ddot{\mathbf{r}} = \frac{I_o}{c} \left[\frac{A}{|\mathbf{r} - \mathbf{r}_j|} \right]^2 \left(\frac{A}{m} \right) \frac{\mathbf{r} - \mathbf{r}_j}{|\mathbf{r} - \mathbf{r}_j|} \dots \quad (2.102)$$

A fraction of the incident radiation is reflected, while the rest is absorbed by the satellite, raising its temperature, and is subsequently re-radiated (Lucas, 1974). This effect depends on the nature of the satellite's surface. A part of the reflected radiation is reflected specularly (angle of incidence = angle of reflection) and the remaining is reflected diffusely. This again depends on the outer surface of the satellite, which can be coated with special materials to control its reflectivity. By Newton's third law the additional force arising from the reflected radiation is opposite to its direction. So the shape of the satellite determines the resultant force. It is obvious from all this that the solar radiation acceleration, like drag, cannot be accurately modelled and, unlike drag, it affects satellites at all altitudes. As in the previous section, the area-to-mass ratio (A/m) is again an important parameter. This must be minimized if solar radiation effects are to be kept small. In contrast, studies of solar radiation have required

the deployment of large balloon-type satellites like PAGEOS and DASH 2 (Lucas, 1974), with a high value of A/m . This ratio is constant only for spherical satellites. In general, it varies with changes in the satellite orientation relative to the sun. In addition, the intensity of the radiation varies with solar activity, which is itself largely unpredictable.

To take account of all the effects mentioned above, a solar radiation reflectance coefficient, C_R , is introduced as a scaling parameter in (2.102). The satellite IF acceleration vector is now given as

$$\ddot{\mathbf{r}} = C_R \frac{I_o}{c} \left[\frac{A}{|\mathbf{r} - \mathbf{r}_j|} \right]^2 \left(\frac{A}{m} \right) \frac{\mathbf{r} - \mathbf{r}_j}{|\mathbf{r} - \mathbf{r}_j|} \quad (2.103)$$

C_R can be introduced as an unknown in the OD process, thus absorbing any uncertainties in the model. Its value is generally greater than unity because of the additional force arising from the reflected radiation. Equation (2.103) gives the acceleration in a direction away from the sun. For satellites with a non-spherical shape, especially when they are equipped with solar panels, the reflected radiation may result in a force that deviates significantly from this direction. This is the case for the NAVSTAR GPS satellites, where solar radiation accelerations are modelled in two directions (Varnum and Chaffee, 1982).

A satellite in its orbit round the earth is not always in sunlight. It occasionally goes into the earth's shadow (umbra), where the solar radiation pressure is cut off. In addition there is a penumbra effect (see fig 2.X) resulting in the radiation pressure diminishing slowly. In order to determine whether or not

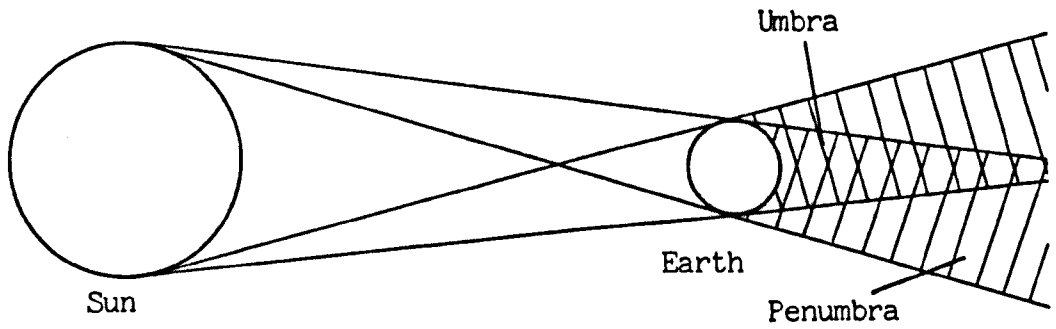


Fig 2.X Umbra and Penumbra

the satellite is in shadow, the rays from the sun are assumed to be parallel. This results in a cylindrical shadow of diameter $2a_m$, where a_m is the earth's equatorial radius modified to account for penumbra effects. The angle, θ , in fig 2.XI is the angle subtended at the centre of the earth (E) between the sun and the satellite (S). This is given by a vector dot product, as

$$\cos \theta = \frac{\mathbf{r} \cdot \mathbf{r}_j}{r r_j} \quad (2.104)$$

If $\cos \theta$ is greater than zero ($\theta < 90^\circ$) then it can be seen from fig 2.XI that the satellite is in sunlight. However, if it is less than zero ($\theta > 90^\circ$), then if

$$\frac{|\mathbf{r} \times \mathbf{r}_j|}{r_j} (= r \sin \theta) < a_m \quad (2.105)$$

the satellite is in shadow.

When the satellite enters the earth's shadow the solar radiation pressure has to be diminished slowly.

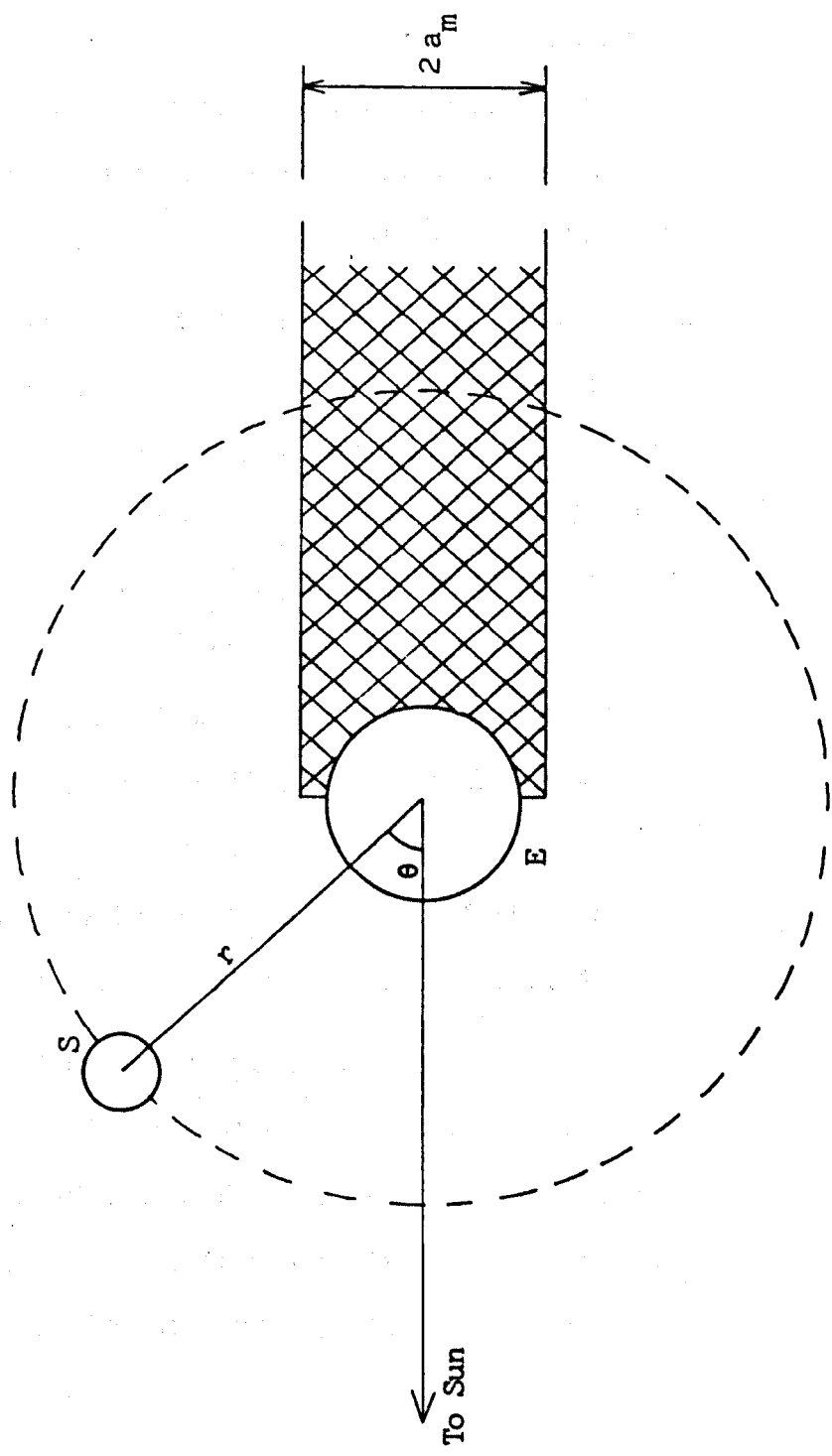


Fig 2.XI Shadow Test for Satellite

There are two reasons for this. In the first place, the reality of the situation is that this radiation does decrease gradually, because of the penumbra. Secondly, the numerical integration procedures (see Section 2.4) that are used to integrate the acceleration and compute the satellite velocity and position as functions of time are not suited to sudden changes in the force model, which can result in the accumulation of large errors.

Albedo radiation is the radiation that is reflected by the earth. This affects the satellite for the same reasons as direct solar radiation, but it decreases, according to the inverse square law, with increasing altitude. In its evaluation the earth is assumed to reflect radiation diffusely and a numerical integration is performed to determine the components of the total albedo flux for the entire earth. The resulting flux has its largest component directed radially outwards, away from the earth. The second component, perpendicular to the first and in the plane defined by the centres of mass of the earth, sun and satellite, is one order of magnitude smaller (Prior, 1972).

Albedo radiation is even less predictable than solar radiation because of temporal changes in the reflective properties of the earth. If its effect is found to be significant, and in the absence of an adequate model, it may be taken into account as an empirical radial acceleration with an adjusted scaling parameter (see Section 2.3.6).

2.3.9 Satellite Thrust

Thrust is available on some satellites in order to enable them to perform a limited number of orbital manoeuvres. The model to account for this acceleration depends on the directions of the jets and on the particular thrust profile (magnitude of thrust as a function of time). The most common case is for along-track, radial and cross-track-accelerations, and these are modelled in the same way as the empirical accelerations in Section 2.3.6. They are scaled according to the known thrust profile for the satellite.

A system called DISCOS (Disturbance Compensation System) is deployed on some satellites, which uses thrust to compensate for all non-gravitational forces (John Hopkins University and Stanford University, 1974). It consists of a proof mass placed in a vacuum inside the satellite and shielded from all surface forces. Any relative displacement of this mass with respect to the walls of the cavity in which it is enclosed is sensed by the satellite and it fires thrusters in order to restore it to its original position. The primary purpose of DISCOS is to eliminate the need to account for the largely unpredictable surface forces, so that the orbit can be determined more accurately. This is even more important when a satellite orbit is predicted for some time in the future, based on the most recent observations. For this reason, DISCOS has found an immediate use in navigation satellites where, for real time navigation, the satellite ephemeris must be predicted as accurately as possible. It is used on the NOVA 1 satellite of the Transit system (Malyevac and Anderle, 1982), where it compensates for all surface forces in the along-track direction.

2.4 NUMERICAL INTEGRATION OF THE EQUATIONS OF MOTION

2.4.1 Equations of Motion

The resultant IF satellite acceleration is the sum of all the components described in the previous section. Some of these may be ignored if they are deemed to be insignificant, depending on the satellite which is being tracked. The equations of motion are 2nd order differential equations, generally given as functions of time, position and velocity, as

$$\ddot{\underline{r}} = \underline{f}(t, \underline{r}, \dot{\underline{r}}) \quad (2.106)$$

where $\ddot{\underline{r}}$: (\ddot{x} , \ddot{y} , \ddot{z}), the resultant IF acceleration

$\dot{\underline{r}}$: (\dot{x} , \dot{y} , \dot{z}), the satellite IF velocity vector

\underline{r} : (x , y , z), the satellite IF position vector

t : UTC time

Orbit integration consists of evaluating the satellite position and velocity vectors as functions of time, $\underline{r}(t)$ and $\dot{\underline{r}}(t)$, using known initial values, \underline{r}_0 and $\dot{\underline{r}}_0$, at initial epoch t_0 . These are given by

$$\dot{\underline{r}}(t) = \dot{\underline{r}}_0 + \int_{t_0}^t \ddot{\underline{r}} dt \quad (2.107)$$

$$\underline{r}(t) = \underline{r}_0 + \int_{t_0}^t \dot{\underline{r}} dt \quad (2.108)$$

The above integrations are performed numerically using the methods described in Section 2.4.2. The satellite ephemeris is computed at discrete time intervals, according to the integration step size, h . The required position and velocity vectors can then be interpolated to the exact appropriate epoch (of the satellite observation) through

some interpolation scheme (see Appendix E). The satellite state vector $(\underline{r}_0, \dot{\underline{r}}_0)$ need not be accurately known, as it can be determined in the least squares adjustment (see Section 2.5).

2.4.2 Numerical Integration of Differential Equations

To simplify the discussion here, a variable, s , is considered which is given by the first order differential equation

$$\dot{s} = \frac{ds}{dt} = f(t, s) \quad (2.109)$$

where $f(t, s)$ denotes some function of time, t , and the variable, s . Successive values of s can be obtained in the same way as in equations (2.107) and (2.108), by

$$s(t) = s_0 + \int_{t_0}^t f(t, s) dt \quad (2.110)$$

where

$$s_0 = s(t_0), \quad (2.111)$$

the initial value of s at time t_0 .

There are two types of methods which can evaluate (2.110) numerically: the first type, known as single-step methods, obtain the next value of the variable in one step, by making use of the current value. The $(i+1)^{th}$ value is given by

$$s_{i+1} = s_i + \int_{t_i}^{t_i+h} f(t, s) dt \quad (2.112)$$

where

$$s_i = s(t_i) = s(t_0 + ih) \quad (2.113)$$

and h : integration step size

A single-step method evaluates the integral

$$F = \int_{t_i}^{t_i+h} f(t,s) dt \quad (2.114)$$

by only making use of s_i . So, starting from s_0 , values of s_1, s_2 , etc., can be computed in successive applications of the method. The second type is the iterative methods, a special case of which are the predictor-corrector methods. In an n^{th} order predictor-corrector scheme the predicted $(i+1)^{\text{th}}$ value of the variable, s_{i+1}^* , is obtained by some numerical formula making use of the previous $n+1$ values. This is now used, along with the previous n values, to evaluate a corrected value, s_{i+1}^{**} . If the difference between s_{i+1}^* and s_{i+1}^{**} exceeds a certain limit, the corrector formula is re-applied using the most recently available value of s_{i+1} . Before the predictor-corrector formulae can come into operation, the first $n+1$ values of s must be available. These can be computed by using a single-step method, starting from the known value of s at t_0, s_0 .

A common family of single-step methods, known as Runge-Kutta, have the form

$$s_{i+1} = s_i + \alpha_1 K_1 + \alpha_2 K_2 + \dots + \alpha_m K_m \quad (2.115)$$

where α : Runge-Kutta coefficients

m : order of the method

The second order Runge-Kutta formulae are given by

$$s_{i+1} = s_i + \alpha_1 K_1 + \alpha_2 K_2 \quad (2.116)$$

$$\text{with } K_1 = hf(t_i, s_i) \quad (2.117)$$

$$K_2 = hf(t_i + \beta h, s_i + \gamma K_1) \quad (2.118)$$

The Taylor series expansion of s_{i+1} about s_i gives:

$$s_{i+1} = s_i + hf_i + \frac{h^2}{2} \left[\frac{\partial f}{\partial t}(t_i, s_i) + f_i \frac{\partial f}{\partial s}(t_i, s_i) \right] + O(h^3) \quad \dots (2.119)$$

where

$$f_i = f(t_i, s_i) \quad (2.120)$$

Expanding K_2 in terms of a Taylor series gives

$$K_2 = h \left[f_i + \beta h \frac{\partial f}{\partial t}(t_i, s_i) + \gamma h f_i \frac{\partial f}{\partial s}(t_i, s_i) \right] + O(h^3) \quad \dots (2.121)$$

Substituting (2.121) in (2.116) leads to:

$$s_{i+1} = s_i + (\alpha_1 + \alpha_2) hf_i + \alpha_2 \beta h^2 \frac{\partial f}{\partial t}(t_i, s_i) + \alpha_2 \gamma h^2 f_i \frac{\partial f}{\partial s}(t_i, s_i) + O(h^3) \quad (2.122)$$

Equating coefficients for (2.119) and (2.122) gives three equations in four unknowns:

$$\begin{aligned} \alpha_1 + \alpha_2 &= 1 \\ \alpha_2 \beta &= \frac{1}{2} \\ \alpha_2 \gamma &= \frac{1}{2} \end{aligned} \quad (2.123)$$

Assigning a value for one of these unknowns determines the other three. A common choice is $\beta = 1$, giving the second order Runge-Kutta formulae as

$$s_{i+1} = s_i + \frac{1}{2} K_1 + \frac{1}{2} K_2 \quad (2.124)$$

with

$$\begin{aligned}
 K_1 &= hf(t_i, s_i) \\
 K_2 &= hf(t_i + h, s_i + K_1)
 \end{aligned}
 \tag{2.125}$$

It was seen that these formulae have been made to agree with the Taylor series expansion of (2.119) to order $O(h^2)$, and terms of $O(h^3)$ and higher orders have been ignored. Similarly, there are 3rd and 4th order Runge-Kutta formulae, which agree with the Taylor series to $O(h^3)$ and $O(h^4)$ respectively. The standard set of 4th order formulae is

$$s_{i+1} = s_i + \frac{1}{6} (K_1 + 2K_2 + 2K_3 + K_4) \tag{2.126}$$

with

$$\begin{aligned}
 K_1 &= hf(t_i, s_i) \\
 K_2 &= hf\left(t_i + \frac{h}{2}, s_i + \frac{K_1}{2}\right) \\
 K_3 &= hf\left(t_i + \frac{h}{2}, s_i + \frac{K_2}{2}\right) \\
 K_4 &= hf(t_i + h, s_i + K_3)
 \end{aligned}
 \tag{2.127}$$

The manner in which these are applied to the OD problem is described in Appendix D.1.

A disadvantage of the Runge-Kutta methods is that there is no way in which the error associated with each step can be evaluated. An estimate of the cumulative error can be obtained if it is assumed that this is proportional to h^m , where m is the order of the method. A value of $s_i^{(1)}$, is obtained after i computations with a step size of h . The true value of s_i is then given by

$$s_i = s_i^{(1)} - Ah^m \tag{2.128}$$

where A : a constant

Another value, $s_i^{(2)}$, is computed after $i/2$ calculations with a step size of $2h$. The true value is now given as

$$s_i = s_i^{(2)} - A (2h)^m \quad (2.129)$$

Combining (2.128) and (2.129) gives the error for the former as

$$\epsilon = Ah^m = \frac{s_i^{(2)} - s_i^{(1)}}{2^m - 1} \quad (2.130)$$

For the 4th order method the error, ϵ , is

$$\epsilon = \frac{1}{15} (s_i^{(2)} - s_i^{(1)}) \quad (2.131)$$

So, for an error estimate to be evaluated, the numerical integration has to be repeated, after halving the step size, until equation (2.130) gives values lying within acceptable bounds. The fact that only one value, s_i , is used in order to generate the next one, s_{i+1} , means that, for sufficient accuracy, the step size of the Runge-Kutta methods has to be kept small. Also, equations (2.127) show that for the 4th order method there are four, usually time-consuming function evaluations per step. For these reasons the Runge-Kutta methods are used only until sufficient values have been calculated, after which a predictor-corrector method takes over.

Predictor-corrector methods replace the function $f(t,s)$ in (2.112) by an interpolation or extrapolation formula which is integrated instead. The Adams-Bashforth method (Spencer et al, 1977) uses the Newton backward difference formula, given by

$$f(t,s) = f_i + p \nabla f_i + \frac{1}{2!} p(p+1) \nabla^2 f_i + \frac{1}{3!} p(p+1)(p+2) \nabla^3 f_i + \dots \quad (2.132)$$

where

$$p = \frac{t - t_i}{h} \quad (2.133)$$

f_i : as given in (2.120)

and ∇^n : the n th backward difference, where, for example:

$$\nabla f_i = f_i - f_{i+1} \quad (2.134)$$

$$\text{and } \nabla^2 f_i = \nabla f_i - \nabla f_{i-1} = f_i - 2f_{i-1} + f_{i-2} \quad (2.135)$$

Substituting (2.132) into (2.112) and integrating from $p=0$ to $p=1$ gives the predicted value

$$s_{i+1}^* = s_i + h \left(f_i + \frac{1}{2} \nabla f_i + \frac{5}{12} \nabla^2 f_i + \frac{3}{8} \nabla^3 f_i + \dots \right) \quad (2.136)$$

The predicted value of f_{i+1} ,

$$f_{i+1}^* = f(t_{i+1}, s_{i+1}^*) \quad (2.137)$$

is now evaluated. The corrector formula is obtained by replacing $f(t,s)$ in equation (2.112) with

$$\begin{aligned} f(t,s) = & f_{i+1}^* + q \nabla f_{i+1}^* + \frac{1}{2!} q(q+1) \nabla^2 f_{i+1}^* + \\ & \frac{1}{3!} q(q+1)(q+2) \nabla^3 f_{i+1}^* + \dots \end{aligned} \quad (2.138)$$

where

$$q = \frac{t - t_{i+1}}{h} \quad (2.139)$$

Integrating from $q = -1$ to $q=0$ gives the corrected value as

$$s_{i+1}^{**} = s_i + h \left(f_{i+1}^* - \frac{1}{2} \nabla f_{i+1}^* - \frac{1}{12} \nabla^2 f_{i+1}^* - \frac{1}{24} \nabla^3 f_{i+1}^* - \dots \right) \quad (2.140)$$

In practice, equations (2.136) and (2.140) are truncated to some order, n , which is the order of the last backward difference to be taken into account. The second order predictor formula, for example, will have $\frac{5}{12} v^2 f_i$ as its last term. This requires that the values f_i , f_{i-1} and f_{i-2} be available. The first three values, needed before this formula is applied, can be obtained by a single-step method such as the 4th order Runge-Kutta, from the known initial value of s at t_0, s_0 .

The error involved in the truncation of the Adams-Bashforth formulae is proportional to the coefficient of the next backward difference. For the second order predictor formula the error is

$$\epsilon^* = s_{i+1}^* - s_{i+1} = \frac{3}{8} C \quad (2.141)$$

where C : a constant

The error in the corresponding corrector is

$$\epsilon^{**} = s_{i+1}^{**} - s_{i+1} = -\frac{1}{24} C \quad (2.142)$$

Eliminating C between (2.141) and (2.142) gives

$$\epsilon^{**} = \frac{1}{10} (s_{i+1}^{**} - s_{i+1}^*) \quad (2.143)$$

In general, the error involved in a truncation after the n^{th} backward difference is

$$\epsilon^{**} = \frac{c_{n+1}}{b_{n+1} - c_{n+1}} (s_{i+1}^{**} - s_{i+1}^*) \quad (2.144)$$

where b_{n+1} : coefficient of the $(n+1)^{\text{th}}$ backward difference from the predictor formula ((2.136))

c_{n+1} : corresponding coefficient from the
corrector formula ((2.140))

From (2.144) it can be seen that an error estimate can be calculated at each step from the difference between the predicted and corrected values. The disadvantage of the Adams-Bashforth method is that the step size, h , cannot be altered if this error estimate is too high (insufficient accuracy) or too low (too many computations). In the former case, the corrector formula can be applied repeatedly, with the most recently available values, until successive values converge. If convergence is slow the order of the formulae can be increased, provided sufficient terms have been evaluated. But if the error estimate is too small, the only efficient method to reduce the number of computations is to alter (increase) the step size. This can only be done if the single-step procedure is re-invoked. However, as can be seen from equations (2.127), the 4th order Runge-Kutta requires 4 function evaluations per step, compared to two for the Adams-Bashforth method. In addition, the step size of the single-step method is, in practice, a fraction of that for the predictor-corrector, resulting in even more computations which may outweigh any subsequent saving from the increased step size of the latter method. For all these reasons it is important to choose carefully, in advance, the step size and order of the Adams-Bashforth formulae. The full formulae, as applied to orbit determination, are given in Appendix D.2.

2.5 LEAST SQUARES ADJUSTMENT AND PARTIAL DERIVATIVES

2.5.1 The Satellite Observations

For a precise orbit determination, the satellite must be observed by a network of globally distributed tracking stations. The requirements of this network are peculiar to the individual satellite, but it is desirable that it should be observed continuously throughout its orbit. There are various types of observations which can be used in OD. The most important for precise work are range, range rate and range difference. Optical observations of the satellite against the background of stars are also used, but for less accurate work.

Range observations can be made in a number of ways. The most common is for a pulse of electromagnetic radiation, either visible or Radio Frequency (RF), to be bounced off the satellite. The time interval between transmission and reception at the tracking station is a measure of the distance travelled or two-way range. This principle is used in Satellite Laser Ranging (SLR) (Ashkenazi, Agrotis and Moore, 1984), where pulses of laser light are bounced off a number of satellites which are equipped with corner-cube retroreflectors. It will also be used in the proposed NAVSAT satellites (Diederich et al, 1984) which will carry transponders capable of relaying RF signals. This technique generally requires complex tracking facilities, but relatively simple equipment on board the satellite. In a variation of the method, the two-way range measurements are made by the satellite, which subsequently transmits them to a central processing facility on earth. The operation of the tracking stations is thus made simpler, but the

satellite has to carry far more complex equipment.

Another commonly used ranging technique is for the satellite to transmit signals which are modulated with coded timing information about their transmission. The difference between this time and the time of reception is a measure of the range. However, this range is corrupted by errors in the satellite and tracking receiver clocks, which have to be calibrated or determined in the OD process. A range measured in this way is termed a pseudo-range and is the principal type of observation for the NAVSTAR GPS satellites (see Chapter 5).

Range rate, that is the rate of change of the range with respect to time, is a function of the relative velocity between satellite and tracking site. It can be observed, if the satellite transmits signals on a stable, continuous radio frequency, by measurements of the Doppler shift of the received frequency. If the Doppler frequency (the difference between received and transmitted frequencies) is integrated between times t and $t + \Delta t$, the number of cycles, or Doppler count, is obtained. This is a measure of the range difference between the two times. Range rate and difference observations are routinely used for tracking the Transit satellites (O'Toole, 1976).

The tracking observations have to be corrected for various system and propagation effects before they can be used in the OD process. The system effects are determined by calibrating against other systems or techniques. Propagation effects arise because of variations in the speed of electromagnetic waves as they enter the ionosphere and troposphere, which cause delays in the arrival of

these signals at the tracking site. The observations have to be corrected to whatever their values would have been, had the signals travelled entirely in vacuo. Ionospheric refraction, which does not affect visible frequencies, is frequency dependent and can be corrected by using dual frequency measurements. Tropospheric refraction is separated into 'dry' and 'wet' components and can be modelled by using surface measurements of atmospheric pressure, temperature and relative humidity, at the time of observation. There are various available tropospheric correction models, the most famous ones being the Hopfield model used for RF observations (Hopfield, 1977), and the Marini and Murray model used for visible frequencies (Marini and Murray, 1973). A more detailed account of ionospheric and tropospheric corrections, as applied to the GPS satellites, is given in Chapter 5.

The technique of least squares adjustment, to determine the most probable values of the various unknown parameters using the satellite tracking data, is outlined in Section 2.5.2. The derivation of the OD observation equations for range observations is treated in Section 2.5.3.

2.5.2 Least Squares Adjustment

The observations taken in order to determine a number, k , of unknown parameters are given as functions of these parameters, as

$$q_i = f_i(x_1, x_2, \dots, x_k) \quad (2.145)$$

where q_i : i^{th} observation (eg range, in OD)

x_1, x_2, \dots, x_k : unknown parameters (eg tracking station coordinates)

The true values of the observations are given by the true values of the unknowns, as

$$\bar{q}_i = f_i (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k) \quad (2.146)$$

If approximate values of the unknowns are assumed, the computed observations will be

$$q_i^c = f_i (x_1^c, x_2^c, \dots, x_k^c) \quad (2.147)$$

and the true observations become

$$\bar{q}_i = f_i (x_1^c + \Delta x_1, x_2^c + \Delta x_2, \dots, x_k^c + \Delta x_k) \quad (2.148)$$

where x^c, x^c, \dots, x_k^c : approximate (computed) values of the unknowns

$\Delta x_1, \Delta x_2, \dots, \Delta x_k$: corrections to the computed values, to give the true unknowns

However, in practice the observations are subject to random errors, and the true observations are replaced by:

$$\bar{q}_i = q_i^o + v_i \quad (2.149)$$

where q_i^o : observed value of observation
 v_i : residual term

The Taylor series expansion of (2.148), ignoring second order terms, is

$$\bar{q}_i = q_i^c + \frac{\partial q_i^c}{\partial x_1} \Delta x_1 + \frac{\partial q_i^c}{\partial x_2} \Delta x_2 + \dots + \frac{\partial q_i^c}{\partial x_k} \Delta x_k \dots \quad (2.150)$$

where

$$\frac{\partial q_i^c}{\partial x_j} = \frac{\partial f_i}{\partial x_j} (x_1^c, x_2^c, \dots, x_k^c) \quad (2.151)$$

Substituting (2.147) and (2.149) in (2.150) gives the least squares observation equation, linearized about the approximate values, as :

$$\frac{\partial q_i^C}{\partial x_1} \Delta x_1 + \frac{\partial q_i^C}{\partial x_2} \Delta x_2 + \dots + \frac{\partial q_i^C}{\partial x_k} \Delta x_k = (q_i^O - q_i^C) + v_i$$

... (2.152)

The partial derivatives, $\frac{\partial q_i^C}{\partial x_j}$, are known as the observation equation coefficients. A set of n observation equations, similar to (2.152), can be expressed in matrix form as

$$Ax = b + v \quad (2.153)$$

where

- A : $(n \times k)$ matrix of observation equation coefficients
- x : $(n \times 1)$ vector of the unknown corrections to the approximate values, $x_1^C, x_2^C, \dots, x_k^C$
- b : $(n \times 1)$ vector of the observed minus computed values of the observations $(q_i^O - q_i^C)$
- v : $(n \times 1)$ vector of residual terms, v_i

These observation equations can be weighted by dividing each one in turn by the estimated (a priori) standard error of the corresponding observation. The purpose of weighting is twofold: in the first place it converts all the terms to dimensionless quantities, making it possible to mix different types of observations. Secondly, it provides a method of attaching more importance to the more accurately observed values. It is also possible to hold certain quantities fixed, by introducing observations with very high weights (low standard errors). There is no need to apply weighting if all observations are of the same type and have identical accuracies. The

weighted observations are given in matrix form, by :

$$W^{\frac{1}{2}} Ax = W^{\frac{1}{2}} b + W^{\frac{1}{2}} v \quad (2.154)$$

where

$$W^{\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & . & . & . & 0 \\ 0 & \frac{1}{\sigma_2} & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & \\ 0 & . & . & . & 0 & \frac{1}{\sigma_n} \end{bmatrix} \quad (2.155)$$

and σ_i : a priori standard error of i^{th} observation.

The least squares solution of (2.154) is that which minimizes the sum of the squares of the weighted residuals. In other words

$$v^T W v = \text{minimum} \quad (2.156)$$

The values of the unknowns which satisfy (2.156) are given by (see Appendix F.1 for derivation)

$$(A^T W A) x = A^T W b \quad (2.157)$$

This can be expressed as

$$N x = d \quad (2.158)$$

where

N : $(k \times k)$ normal matrix $(A^T W A)$

d : $(k \times 1)$ right hand side sector $(A^T W b)$

Equations (2.158) are known as the normal equations, and can be solved by a number of methods to give the vector of unknowns, x (Ashkenazi, 1967 and 1969). The matrix, N , is symmetric and positive-definite (see Appendix F.2) and an appropriate method is Choleski's method of symmetric decomposition described in Appendix F.3.

The covariance matrix of the vector of unknowns is defined as :

$$\sigma_{xx} = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \dots & \sigma_{x_1 x_k} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \dots & \sigma_{x_2 x_k} \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{x_k x_1} & \dots & \dots & \sigma_{x_k}^2 \end{bmatrix} \quad (2.159)$$

where $\sigma_{x_i}^2$: variance of the i^{th} unknown, x_i
 $\sigma_{x_i x_j}$: covariance of the pair of elements, x_i
 and x_j

This matrix is given by (Ashkenazi, 1970)

$$\sigma_{xx} = \sigma_o^2 N^{-1} \quad (2.160)$$

where σ_o^2 : unbiased estimate of the variance of an observation of unit weight, computed from

$$\sigma_o^2 = \frac{v^T W v}{n - k} \quad (2.161)$$

A derivation of this formula can be found in Ashkenzi (1970). The value of the unit variance, σ_0 , is an indication of whether the a priori standard errors have been estimated correctly and, in theory, it should take a value of unity. If however no weighting has been applied, σ_0 is an indication of the magnitude of the root-mean-square (rms) residual of the adjusted observations.

The covariance matrix, apart from giving the variances and covariances of the unknowns, can also be used to compute the a posteriori (after adjustment) variances of the observations and of other important quantities. This can be done by making use of Gauss's propagation of error theorems, which give the variance of the i^{th} observation as

$$\sigma_{qi}^2 = \left(\frac{\partial q_i}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial q_i}{\partial x_2}\right)^2 \sigma_{x_2}^2 + \dots + \left(\frac{\partial q_i}{\partial x_k}\right)^2 \sigma_{x_k}^2 + 2\left(\frac{\partial q_i}{\partial x_1} \frac{\partial q_i}{\partial x_2}\right) \sigma_{x_1 x_2} + \dots$$

... (2.162)

and the covariance of observations i and j , as

$$\sigma_{qi qj} = \left(\frac{\partial q_i}{\partial x_1} \frac{\partial q_j}{\partial x_1}\right) \sigma_{x_1}^2 + \left(\frac{\partial q_i}{\partial x_2} \frac{\partial q_j}{\partial x_2}\right) \sigma_{x_2}^2 + \dots + \left(\frac{\partial q_i}{\partial x_1} \frac{\partial q_j}{\partial x_2} + \frac{\partial q_i}{\partial x_2} \frac{\partial q_j}{\partial x_1}\right) \sigma_{x_1 x_2} + \dots$$

... (2.163)

From (2.162) and (2.163) it follows that, for

$$\hat{b} = Ax \quad (2.164)$$

where \hat{b} : ($n \times 1$) vector of a posteriori observed minus computed values

then

$$\sigma_{\hat{b}\hat{b}} = A \sigma_{xx} A^T \quad (2.165)$$

and σ_{bb} is the covariance matrix of the adjusted observations. It should be noted here that the covariance matrix of the observations is the same as that of the observed minus computed values. The variances and covariances of all other quantities which are functions of the unknowns, can be obtained by deriving expressions similar to (2.162) and (2.163).

2.5.3 The Orbit Determination Observation Equations

It was explained in Section 2.5.1 that the satellite can be tracked using a number of observation types. The equations that will be presented here are only for range observations, but a similar treatment can be easily applied to other types, such as range rate.

A range between station s and the satellite is independent of reference frame and is given by

$$l = [(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2]^{\frac{1}{2}} \quad (2.166)$$

or by

$$l = [(X - X_s)^2 + (Y - Y_s)^2 + (Z - Z_s)^2]^{\frac{1}{2}} \quad (2.167)$$

where

\underline{r} , \underline{R} : (x, y, z) and (X, Y, Z) , the satellite IF and EF coordinates respectively

\underline{r}_s , \underline{R}_s : (x_s, y_s, z_s) and (X_s, Y_s, Z_s) , the IF and EF coordinates of station s .

All these position vectors must be known at the time of observation. This is the instant at which the signal reaches (or leaves) the satellite, for satellite coordinates, and the instant at which it

leaves (or arrives at) the station, for station coordinates. The satellite coordinates are given at discrete intervals according to the numerical integration time step (see Section 2.4), and can be interpolated using a suitable interpolation scheme (see Appendix E.2). The full least squares observation equation for an observed range, linearized about some approximate values, is (see Section 2.5.2).

$$\sum_{i=1}^6 \frac{\partial l}{\partial r_i^o} \Delta r_i^o + \sum_{i=1}^{np} \frac{\partial l}{\partial p_i} \Delta p_i + \sum_{i=1}^{ne} \frac{\partial l}{\partial u_i} \Delta u_i + ,$$

$$\sum_{i=1}^3 \frac{\partial l}{\partial R_{si}} \Delta R_{si} = (l^o - l^c) + v \quad (2.168)$$

where r_i^o : one of the six IF components of the satellite state vector ($\underline{r}_o, \dot{\underline{r}}_o$)

p_i : one of the np force model unknowns to be included in the solution (eg $C_R, C_D, C_a, \bar{C}_n^m, \bar{S}_n^m, GM$, as defined in Section 2.3)

u_i : one of the ne earth rotation parameters (ERP's), which could be x_p, y_p or UT1-UTC (see Section 2.2.3)

R_{si} : one of the components of the EF position vector, \underline{R}_s , of station s

l^o, l^c : values of observed and computed range respectively

The actual observation equations do not need to include all the terms mentioned above, and any quantities which are known in advance can be excluded. In addition, it is possible to constrain a number of unknowns by introducing observation equations of the form

$$\Delta q = (q^o - q^c) + v \quad (2.169)$$

where q : any one of the unknowns in (2.168)

q^o : observed (required) value of the unknown

q^c : computed value which is usually the same as q^o

and attributing to them appropriate weights. The observations from all the ns tracking stations are assembled, observation equations are formed and then solved using least squares, as described in Section 2.5.2, for the required unknown corrections Δr_i^o , Δp_i , Δu_i and ΔR_{si} . For the formation of the observation equations, the coefficients $\frac{\partial l}{\partial r_i^o}$, $\frac{\partial l}{\partial p_i}$, $\frac{\partial l}{\partial u_i}$ and $\frac{\partial l}{\partial R_{si}}$ need to be evaluated. The rest of this section will deal with the computation of the observation equation coefficients for each type of unknown. The most straightforward are those for tracking station coordinates, and they are given by differentiating equation (2.167), as

$$\frac{\partial l}{\partial R_{si}} = \frac{R_{si} - R_i}{l} \quad (2.170)$$

where R_i : one of the EF components of the satellite position vector, \underline{R}

The coefficient for an ERP unknown, assuming that these only affect station (and not satellite) coordinates, is

$$\frac{\partial l}{\partial u_i} = \sum_{j=1}^3 \frac{\partial l}{\partial R_{sj}} \frac{\partial R_{sj}}{\partial u_i} \quad (2.171)$$

The derivatives $\frac{\partial l}{\partial R_{sj}}$ are computed as in (2.170). The partials $\frac{\partial R_{sj}}{\partial u_i}$ are given in vector form, as

$$\frac{\partial \underline{R}_S}{\partial u_i} = \left(\frac{\partial X_S}{\partial u_i}, \frac{\partial Y_S}{\partial u_i}, \frac{\partial Z_S}{\partial u_i} \right) \quad (2.172)$$

But, from (2.28)

$$\underline{R}_S = P E N Q \underline{r}_S$$

and

$$\frac{\partial \underline{R}_S}{\partial u_i} = \frac{\partial}{\partial u_i} (P E N Q \underline{r}_S) \quad (2.174)$$

Taking the polar motion component, x_p , as an example,

$$\frac{\partial \underline{R}_S}{\partial x_p} = \frac{\partial P}{\partial x_p} \cdot E N Q \underline{r}_S \quad (2.175)$$

where

$$\frac{\partial P}{\partial x_p} = \frac{\partial}{\partial x_p} [R_2(-x_p)] \cdot R_1(-y_p) \quad (2.176)$$

and

$$\frac{\partial}{\partial x_p} [R_2(-x_p)] = \begin{bmatrix} -\sin x_p & 0 & \cos x_p \\ 0 & 0 & 0 \\ -\cos x_p & 0 & -\sin x_p \end{bmatrix} \quad (2.177)$$

The observation equation coefficient for y_p is derived similarly.

When UT1 - UTC is being determined, the observation equations are set up in terms of a correction, $\Delta GAST$, to Greenwich Apparent

Sidereal Time (GAST). The required coefficient, $\frac{\partial l}{\partial GAST}$, is computed from (2.171), using

$$\frac{\partial \vec{R}_S}{\partial \text{GAST}} = P \cdot \frac{\partial E}{\partial \text{GAST}} \cdot NQ \vec{r}_S \quad (2.178)$$

where E is the earth rotation matrix (see Section 2.2.3).

The correction, ΔGAST , is subsequently converted to its Universal Time equivalent, ΔUT1 ($= \Delta (\text{UT1} - \text{UTC})$), by making use of (2.22).

The coefficient for a satellite state vector unknown is

$$\frac{\partial l}{\partial r_i^0} = \sum_{j=1}^3 \frac{\partial l}{\partial r_j} \frac{\partial r_j}{\partial r_i^0} \quad (2.179)$$

and for a force model unknown,

$$\frac{\partial l}{\partial p_i} = \sum_{j=1}^3 \frac{\partial l}{\partial r_j} \frac{\partial r_j}{\partial p_i} \quad (2.180)$$

For example

$$\frac{\partial l}{\partial x_0} = \frac{\partial l}{\partial x} \frac{\partial x}{\partial x_0} + \frac{\partial l}{\partial y} \frac{\partial y}{\partial x_0} + \frac{\partial l}{\partial z} \frac{\partial z}{\partial x_0} \quad (2.181)$$

The partial derivatives $\frac{\partial l}{\partial r_j}$ in equations (2.179) and (2.180) are easily evaluated by differentiating (2.166), as

$$\frac{\partial l}{\partial r_j} = \frac{r_j - r_{sj}}{1} \quad (2.182)$$

but the derivatives of satellite position with respect to the state vector and force model unknowns, $\frac{\partial r_j}{\partial r_i^0}$ and $\frac{\partial r_j}{\partial p_i}$, have to be obtained by numerical integration. The acceleration partials are integrated once to give the velocity partials, and twice to give the position partials as functions of time, as

$$\frac{\partial \dot{r}_j}{\partial r_i^0} = \left[\frac{\partial \dot{r}_j}{\partial r_i^0} \right]_{t_0} + \int_{t_0}^t \frac{\partial \ddot{r}_j}{\partial r_i^0} dt \quad (2.183)$$

$$\frac{\partial \mathbf{r}_j}{\partial \mathbf{r}_i^0} = \left[\frac{\partial \mathbf{r}_j}{\partial \mathbf{r}_i^0} \right]_{t_0} + \int_{t_0}^t \frac{\partial \dot{\mathbf{r}}_j}{\partial \mathbf{r}_i^0} dt \quad (2.184)$$

with similar expressions for $\frac{\partial \mathbf{r}_j}{\partial \mathbf{p}_i}$. The acceleration derivatives with respect to the state vector unknowns are

$$\frac{\partial \ddot{\mathbf{r}}_j}{\partial \mathbf{r}_i^0} = \sum_{k=1}^3 \frac{\partial \ddot{\mathbf{r}}_j}{\partial \mathbf{r}_k} \frac{\partial \mathbf{r}_k}{\partial \mathbf{r}_i^0} + \sum_{k=1}^3 \frac{\partial \ddot{\mathbf{r}}_j}{\partial \dot{\mathbf{r}}_k} \frac{\partial \dot{\mathbf{r}}_k}{\partial \mathbf{r}_i^0} \quad (2.185)$$

and those with respect to the force model parameters are :

$$\frac{\partial \ddot{\mathbf{r}}_j}{\partial \mathbf{p}_i} = \sum_{k=1}^3 \frac{\partial \ddot{\mathbf{r}}_j}{\partial \mathbf{r}_k} \frac{\partial \mathbf{r}_k}{\partial \mathbf{p}_i} + \sum_{k=1}^3 \frac{\partial \ddot{\mathbf{r}}_j}{\partial \dot{\mathbf{r}}_k} \frac{\partial \dot{\mathbf{r}}_k}{\partial \mathbf{p}_i} + \frac{\partial \ddot{\mathbf{r}}_j^*}{\partial \mathbf{p}_i} \quad (2.186)$$

where $\dot{\mathbf{r}}_k$: one of the components of the IF satellite velocity vector, $\dot{\mathbf{r}}$, at observation epoch

$\ddot{\mathbf{r}}_j$: one of the components of the satellite IF acceleration vector, $\ddot{\mathbf{r}}$, at observation epoch

$\frac{\partial \ddot{\mathbf{r}}_j^*}{\partial \mathbf{p}_i}$: derivative obtained by direct differentiation of the constituent of the force model which contains \mathbf{p}_i (see Appendix C.2)

For example, in

$$\begin{aligned} \frac{\partial \ddot{x}}{\partial C_R} &= \frac{\partial \ddot{x}}{\partial x} \frac{\partial x}{\partial C_R} + \frac{\partial \ddot{x}}{\partial y} \frac{\partial y}{\partial C_R} + \frac{\partial \ddot{x}}{\partial z} \frac{\partial z}{\partial C_R} + \frac{\partial \ddot{x}}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial C_R} \\ &+ \frac{\partial \ddot{x}}{\partial \dot{y}} \frac{\partial \dot{y}}{\partial C_R} + \frac{\partial \ddot{x}}{\partial \dot{z}} \frac{\partial \dot{z}}{\partial C_R} + \frac{\partial \ddot{x}^*}{\partial C_R} \end{aligned} \quad (2.187)$$

the term $\frac{\partial \ddot{x}^*}{\partial C_R}$ is obtained by differentiating equation (2.103),

for the satellite acceleration arising from solar radiation, with respect to C_R . The partial derivatives $\frac{\partial \ddot{r}_j}{\partial r_k}$ and $\frac{\partial \ddot{r}_j}{\partial r_k^*}$ are computed by differentiating the various components of the force model in turn, and summing the corresponding partials for all the components (see Appendix C.2)

Equations (2.185) and (2.186) can be combined and expressed in matrix form, as

$$A_m = D1 \cdot X_m + D2 \cdot V_m + A_f \quad (2.188)$$

where A_m : $(3 \times (6 + np))$ matrix of acceleration partials

$$\frac{\partial \ddot{r}_j}{\partial r_i^0} \quad \text{and} \quad \frac{\partial \ddot{r}_j}{\partial p_i}$$

V_m : $(3 \times (6 + np))$ matrix of velocity partials $\frac{\partial \dot{r}_j}{\partial r_i^0}$ and

$$\frac{\partial \dot{r}_j}{\partial p_i}$$

X_m : $(3 \times (6 + np))$ matrix of position partials $\frac{\partial r_j}{\partial r_i^0}$ and

$$\frac{\partial r_j}{\partial p_i} \quad (\text{see equations (2.179) and (2.180)})$$

A_f : $(3 \times (6 + np))$ matrix of direct derivatives, $\frac{\partial \ddot{r}_j^*}{\partial p_i}$

(see Appendix C.2)

$D1$: (3×3) matrix of $\frac{\partial \ddot{r}_j}{\partial r_k}$

$D2$: (3×3) matrix of $\frac{\partial \ddot{r}_j}{\partial \dot{r}_k}$

In the absence of drag, along-track and cross-track accelerations, the satellite acceleration is not a function of velocity and the

$\frac{\partial \ddot{\mathbf{r}}_j}{\partial \dot{\mathbf{r}}_k}$ are all zero. D_2 is then a zero matrix. The computation of equations (2.188) is explained in more detail in Appendix C.2.

These give the acceleration partials, A_m , as functions of the position and velocity partials, X_m and V_m . There are $18 + 3np$ second order differential equations, and they are similar to (2.106) which gives the satellite acceleration as a function of position and velocity. A numerical integration scheme, as outlined in Section 2.4 and Appendix D, gives the matrices V_m and X_m as functions of time. To start the integration, the matrices of initial values, X_m^0 and V_m^0 , at initial epoch t_0 , are required. These are given by

$$X_m^0 = \begin{bmatrix} 1 & 0 & 0 & . & . & . & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & . & . & . & 0 \\ 0 & 0 & 1 & 0 & . & . & . & . & . & 0 \end{bmatrix} \quad (2.189)$$

and

$$V_m^0 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & . & . & . & . & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & . & . & 0 \end{bmatrix} \quad (2.190)$$

To explain (2.189) and (2.190) take, for example, the first element of the first row of X_m^0 . This is

$$\left[\frac{\partial x}{\partial x_0} \right]_{t_0} = \frac{\partial x_0}{\partial x_0} = 1 \quad (2.191)$$

The second element is

$$\left[\frac{\partial x}{\partial y_o} \right]_{t_o} = \frac{\partial x_o}{\partial y_o} = 0 \quad (2.192)$$

because a change in one component of the state vector (e.g. y_o) does not produce any change in the other components (e.g. x_o).

For the same reason, for a force model unknown

$$\left[\frac{\partial x}{\partial p_i} \right]_{t_o} = \frac{\partial x_o}{\partial p_i} = 0 \quad (2.193)$$

As least squares provides only a first order correction, it may be possible to ignore some of the components of the force model when evaluating D1 and D2. In the initial solution, if the corrections to the approximate values of the unknowns are large, the adjustment is repeated with the most recent values until they converge (i.e. corrections become negligible). The omission of the less significant constituents of the force model when computing the partials should not adversely affect the rate of convergence.

2.5.4 Orbit Determination Adjustment Requirements

When all the required satellite tracking data have been collected, a set of observation equations similar to (2.168) is solved by least squares (see Section 2.5.2), in order to determine the various unknowns described in the previous section. The total number of unknown parameters is

$$n_T = 6 + 3n_s + n_p + n_e \quad (2.194)$$

where

- n_s : number of tracking stations
- n_p : number of force model unknowns
- n_e : number of earth rotation parameters (ERP's)

The set of tracking stations constitutes a three-dimensional network. In any 3-D adjustment, the normal equations (equations (2.158)) are singular unless a number of constraints is imposed. There are seven such constraints : three to locate the origin, three to orientate the coordinate axes, and one to provide the scale for the network. The rank deficiency of the normals is the number of such conditions which are not satisfied.

In orbit determination, the three origin conditions are satisfied by the gravity field. The spherical harmonic coefficients C_1^0 , C_1^1 and S_1^1 represent the first moments of mass of the earth about the origin of coordinates. These coefficients are set to zero in the geopotential expansion of (2.30), thus locating the origin at the earth's mass centre. The scale of the network is provided by the adopted values of the speed of light, c , and of the earth's gravitational constant, GM . These two quantities must be compatible,

and if the speed of light is held fixed the value of GM can be determined in the adjustment. There are now the three orientation deficiencies left. Two of these are satisfied by the polar motion values, x_p and y_p , which locate the Z-axis of the EF reference frame with respect to the earth's instantaneous spin axis. This still leaves the X-axis free to rotate, and it is defined by fixing the longitude of one of the tracking stations. This is achieved by introducing an observation equation, similar to (2.169), given by

$$\Delta\lambda_s = (\lambda_s^O - \lambda_s^C) + v \quad (2.195)$$

where λ_s : EF spheroidal longitude of station s (see Appendix B.3)

From equation (B.18),

$$\lambda_s = \tan^{-1} \frac{Y_s}{X_s} \quad (2.196)$$

The observation equation (2.195) can be expressed in terms of X_s and Y_s , as

$$\frac{\partial\lambda_s}{\partial X_s} \Delta X_s + \frac{\partial\lambda_s}{\partial Y_s} \Delta Y_s = (\lambda_s^O - \lambda_s^C) + v \quad (2.197)$$

where

$$\frac{\partial\lambda_s}{\partial X_s} = \frac{-Y_s}{X_s^2 + Y_s^2} \quad (2.198)$$

$$\frac{\partial\lambda_s}{\partial Y_s} = \frac{X_s}{X_s^2 + Y_s^2} \quad (2.199)$$

and ΔX_s and ΔY_s are the unknown corrections to the provisional values of the X and Y-coordinates of station s. Equation (2.197) is the 'observation' equation that, when given a sufficiently high

weight (small standard error) fixes the longitude of the station to the required value, λ_s^0 .

The seven parameters described above define the EF reference frame. It is clear that since the polar motion values, x_p and y_p , are two of the orientation constraints, it is not possible to solve for all the station coordinates and for polar motion simultaneously (see Section 4.2.4). If x_p and y_p are being determined, then the latitudes of two stations, in addition to the fixed longitude, have to be constrained. In practice it is better to fix a number of tracking stations and use their coordinates in order to define the orientation of the EF reference frame. Initially, these coordinates can be derived by using externally provided (e.g. by the BIH) polar motion values. This process will 'align' the network to the external values, removing the biases that would otherwise exist between the two systems.

It was seen earlier (Section 2.4) that the satellite coordinates are derived in the IF. This is related to the EF reference frame by the rotation matrices described in Section 2.2, and in particular by the earth rotation matrix, E , of equation(2.24). E represents a rotation about the X-axis through GAST. In Section 2.2.3 it was shown that GAST is equivalent to universal time, UT1. This implies that absolute values of UT1-UTC (see Section 2.5.3) cannot be determined at the same time as the satellite state vector, because they would allow the satellite orbit to rotate to any arbitrary 'inertial frame'. The OD process is thus not suited to determining absolute values of UT1-UTC. However, it is possible to compute changes in UT1-UTC (or GAST) from any initial value. The problem now arises of how one would go about determining earth rotation from

satellite tracking data (see Section 4.2.5). One approach is to determine a satellite orbit over a long arc, say one month, using externally provided values of UT1-UTC and an appropriate set of observations. This orbit can then be held fixed in a number of adjustments with smaller portions of data, spanning 1-2 days, in order to determine values of UT1-UTC for the mid-epochs of the new data sets. The disadvantage of this approach is that it is difficult to maintain continuity over successive long arcs.

A better method is to determine the satellite orbit over a number of days, in an adjustment which also incorporates a number of, say daily, UT1-UTC unknowns. The singularity that exists between these and the satellite state vector can be overcome by fixing the first UT1-UTC unknown at any chosen value. All the other unknowns are determinable, because the fixed value determines the relationship between the IF and EF reference frames, and the subsequent solutions for UT1-UTC represent the changes in the earth's rotation that have occurred since the beginning of the data set. Continuity can be maintained in this approach by overlapping successive long arcs, and fixing the first UT1-UTC unknown of the present solution at the value derived from the previous adjustment. Another advantage of this method is that all the unknowns can be determined in a single adjustment rather than the two-stage process of the first approach.

CHAPTER THREE

UNIVERSITY OF NOTTINGHAM ORBIT DETERMINATION

SOFTWARE

3.1 Introduction

Orbit determination software was developed at Nottingham in order to process SLR observations of the Laser Geodynamics Satellite (LAGEOS), and the models used are specifically suited for this satellite. However, with some alterations, the programs can be used in conjunction with other satellites and types of data.

LAGEOS was launched in 1976 in a near-circular orbit, at an altitude of 5900 km and an inclination of 110° to the equator. At this altitude it orbits the earth every 225 minutes. It is a spherical satellite with a mass of 408 kg and a diameter of 60 cm, and its surface is completely covered with corner-cube retro-reflectors. Its low area-to-mass (A/m) ratio minimizes the effects of surface forces and its high altitude eliminates the need to model air drag (see Section 2.3.7). Gravitational forces need to be modelled very accurately, and the software allows for the direct attractions of the moon, sun, Venus, Mars, Jupiter and Saturn (see Section 2.3.3). Lunar and solar solid earth tides (Section 2.3.4), ocean tides (Section 2.3.5), solar radiation (Section 2.3.8), and an empirical along-track acceleration (Section 2.3.6) are also modelled. The latter takes into account any unmodelled drag-like effects. A choice of two geopotential models, GEM10 and GEM-L2 (see Section 2.3.2), truncated to degree and order 20, is available.

The software package was written in FORTRAN 77 for the ICL 2977 computer of the University of Nottingham. It consists of four program modules, ORBIT, CHEBPOL, SOAP and DATPAK, which are illustrated in the flow chart in fig 3.I. Of these, ORBIT and CHEBPOL were developed by the author, SOAP was written jointly with

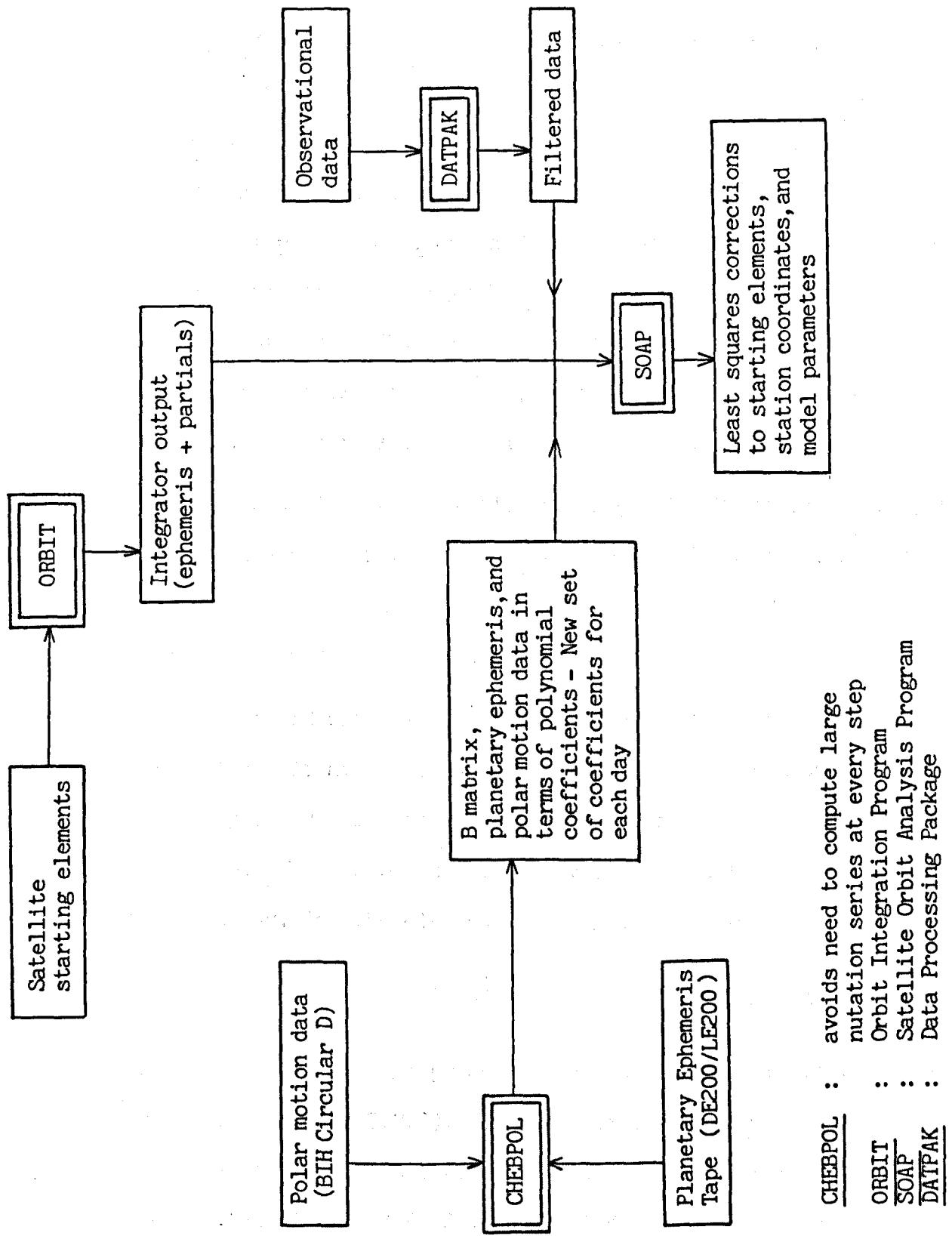


Fig 3.1 Nottingham Orbit Determination Software

Mr. T. Moore of Nottingham University, and DATPAK was developed by Mr. T. Moore with the author's assistance.

ORBIT is the program that performs the numerical integration of the orbit. It takes the satellite state vector and, using an appropriate force model (Section 2.3) and numerical integration procedures (Section 2.4), computes and outputs the satellite ephemeris ($\underline{r}(t)$ and $\dot{\underline{r}}(t)$) and the required observation equation partial derivatives (matrices X_m and V_m in Section 2.5.3) as functions of time. The program was tested by comparing it with the output of a similar program, developed independently by Dr. A. T. Sinclair of the Royal Greenwich Observatory.

The Satellite Orbit Analysis Program (SOAP) takes the output from ORBIT and the pre-processed (filtered) SLR observations and performs the least squares adjustment (see Section 2.5) for a specified number of unknowns.

The third program, CHEBPOL, serves two functions: in the first place, it avoids the need to evaluate a large nutation series for every integration step (in ORBIT) and for every observation (in SOAP). This is because it derives interpolation coefficients for each day, for the elements of matrix B, where

$$B = Q^T N^T \quad (3.1)$$

and Q and N are the precession and nutation matrices respectively (see Section 2.2.2). It also computes interpolation coefficients for the nutation in longitude, $\Delta\phi$, the planetary coordinates of DE200/LE200 (see Section 2.3.3) and the ERP'S ($x_p, y_p, UT1-UTC$). Secondly, it makes this data available to both

ORBIT and SOAP by storing it in a random access file.

This chapter documents the main features of programs CHEBPOL (Section 3.2), ORBIT (Section 3.3) and SOAP (Section 3.4). DATPAK is a pre-processing program, which filters and compresses the SLR data. It will not be described in detail, but its main features are :

- ① A 10th order polynomial is fitted by least squares to each LAGEOS pass, and the rms residual of the fit, σ , is calculated.
- ② All observations with residuals greater than 2σ are rejected.
- ③ If σ is greater than 10 m, steps ① and ② are repeated up to four times.
- ④ The observations are separated into small sets, spanning intervals of one minute, and a 7th order polynomial is fitted to each set using least squares.
- ⑤ If the rms residual of this fit is greater than one metre, the minute of data is rejected.
- ⑥ A point is taken near the middle of each accepted polynomial, which is called the 'normal point'.

These normal points are used as the data input in SOAP.

The strategy outlined above (Ashkenazi, Agrotis and Moore, 1984) eliminates noisy observations and reduces considerably the amount of data that is handled in SOAP. The use of normal points produces solutions which are, to all purposes, identical to those obtained with the full set of contributing observations, at a fraction of the cost.

3.2 CHEBYSHEV POLYNOMIAL PROGRAM (CHEBPOL)

3.2.1 Data Input

CHEBPOL evaluates polynomial coefficients for 30 days, starting from a specified start date. It requires as input :

- (i) seven sets of 5-day x_p , y_p and UT1-UTC values from BIH Circular D (see Section 2.2.3).
- (ii) the heliocentric IF lunar and planetary coordinates from DE200/LE200 (see Section 2.3.3). These are given in Astronomical Units (AU) for $0.^{h0}$ TDB of each day and they are stored on magnetic tape for the years 1979 to 2006. To reduce filestore requirements, a portion of this ephemeris covering the years 1980 - 84 has been transferred to a random access file, and this is used as the input in CHEBPOL.
- (iii) the year and day of year, from which to start evaluating coefficients. This must coincide with the first day for which the ERP's are given.
- (iv) the output mode. This varies the amount of line printer output.

A simplified flow chart of CHEBPOL, showing the different types of input, is given in fig 3.II.

3.2.2 Program Description and Output

CHEBPOL evaluates Chebyshev polynomial coefficients for each day, for :

- (i) the 9 elements of matrix B, defined in equation (3.1) (see Section 2.2.2)

- (ii) the nutation in longitude, $\Delta\psi$ (see equations (2.12)).
This is required in order to convert GMST to GAST, which is used in the computation of the earth rotation matrix, E (see Section 2.2.3).
- (iii) the geocentric IF coordinates of the moon, sun, Venus, Mars, Jupiter and Saturn. These are required by ORBIT to allow for third body gravitational effects and for solar radiation pressure on the satellite (see Sections 2.3.3 and 2.3.8), and by both ORBIT and SOAP to evaluate tidal corrections (Section 2.3.4).

CHEBPOL also evaluates linear interpolation coefficients for the ERP's, from the 5-day BIH values.

The output is on a random access file and spans 30 days from the specified start day. This file is subsequently used as input to both ORBIT (Section 3.3) and SOAP (Section 3.4). Each output record contains all the coefficients for one day. The first number on the record is the Julian day number (JD) of the day in question. This is defined such that

$$2000 \text{ January } 1 \text{ at } 12^{\text{h}}.0 \text{ UT} = \text{JD } 2451545.0 \text{ UT} \quad (3.2)$$

and one Julian day corresponds to one solar day. So, for example,

$$1980 \text{ September } 2, 0^{\text{h}}.0 \text{ UTC} = \text{JD } 2444484.5 \text{ UTC} \quad (3.3)$$

The Julian day number provides an unambiguous and easy way of specifying the date. The interpolation coefficients are referred to $0^{\text{h}}.0 \text{ UTC}$ of the particular day, and for any known UTC epoch, t_{UTC} , and Julian date, JD, all the required values can be computed.

The Chebyshev representation of a function, $f(t)$, of a variable, t , (see Appendix E.1) consists of evaluating coefficients $\alpha_0, \alpha_1, \dots, \alpha_n$, where n is the order of the Chebyshev polynomial. These coefficients are valid for the interval, Δt , for which they are computed.

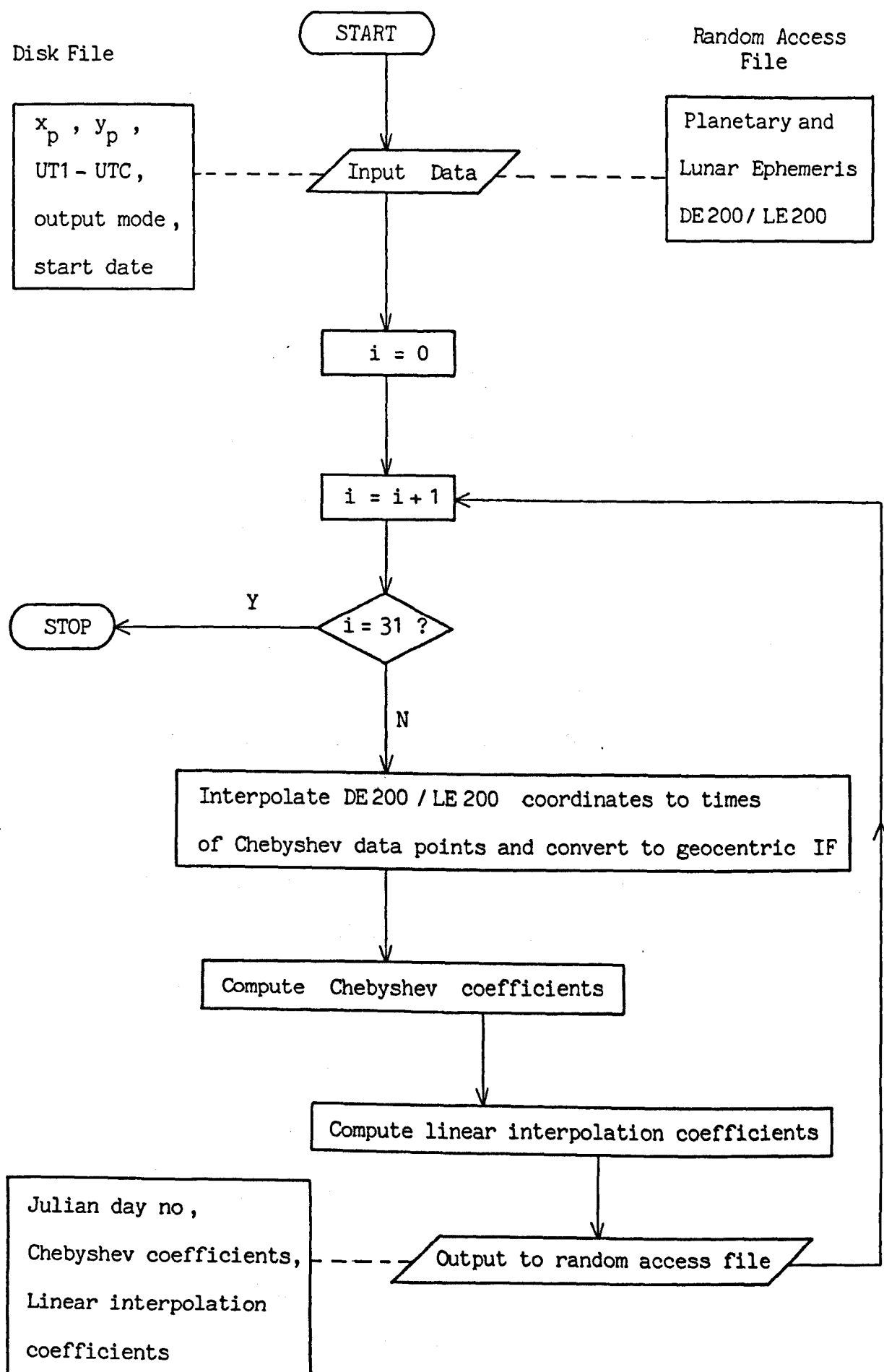
In CHEBPOL, the variable is UTC time since $0^h.0$ of the day in question, t_{UTC} , and the interval, Δt , is $24^h.0$. Any value of a function (eg $\Delta\phi$) within this interval is given by

$$f(t_{UTC}) = \sum_{k=0}^{\ell} \alpha_k \cos k\theta \quad (3.4)$$

where

$$\theta = \cos^{-1} \left[\frac{t_{UTC} - 12^h.0}{24^h.0} \right] \quad (3.5)$$

The Chebyshev coefficients are evaluated for degree $n = 10$, but they are truncated in (3.4) to $\ell = 4$. Five coefficients, $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, and α_4 , for each function, are output in the random access file. Since there are 18 (6×3) planetary (and lunar) coordinates, 9 elements of matrix B , and one value of $\Delta\phi$, the total number of Chebyshev coefficients per day is 140. For the computation of these coefficients, each function has to be evaluated at 11 ($n+1$) data points (see Appendix E.1) within the 24-hour interval. The data points are defined by equations (E.3) and (E.4). They present no problem for the elements of B and for $\Delta\phi$, since the precession and nutation angles (see Section 2.2.2) are given as continuous functions. The planetary and lunar coordinates, however, are given at discrete intervals, at $0^h.0$ TDB of each day. They first have to be interpolated to the UTC epochs that are specified by the Chebyshev data points. This interpolation is performed using Everett's formula

Fig.3.II Flow Diagram for CHEBPOL

(see Appendix E.2) up to the fourth central difference.

The coordinates are subsequently converted to the geocentric IF (see Section 2.3.3), and from Astronomical Units to metres ($1 \text{ AU} = 1.4959787 \times 10^{11} \text{ m}$).

The ERP's are given for any UTC time by linear interpolation. For example

$$x_p = \beta_0 + \beta_1 t_{\text{UTC}} \quad (3.6)$$

and the output of CHEBPOL consists of coefficients β_0 and β_1 for x_p , y_p and $UT1 - \text{UTC}$, for each day. These coefficients are derived by linear interpolation of the 5-day BIH values of Circular D.

A simplified flow chart of CHEBPOL is given in fig 3.II. The main advantage of using this program is that it avoids the need to compute a large nutation series of 106 terms for every integration step (in ORBIT) and for every observation (in SOAP). It also makes the last two programs far less complex, because they do not need to incorporate the precession and nutation theories (Section 2.2.2). For these reasons, CHEBPOL is also being used in the University of Nottingham Very Long Baseline Interferometry (VLBI) software package.

3.3 THE SATELLITE ORBIT INTEGRATION PROGRAM (ORBIT)

3.3.1 Numerical Integration and Force Model Definition

Orbit performs the numerical integration of the IF satellite acceleration, \ddot{r} , (see Section 2.4), and acceleration partial derivatives, A_m (see Section 2.5.3). A total of $[3 + 3 \times (6 + np)]$ second order differential equations (np = number of force model unknowns) are

integrated to give the satellite IF position and velocity vectors, \underline{r} and $\dot{\underline{r}}$, and the position and velocity partials, X_m and V_m , based on the initial elements, \underline{r}_0 , $\dot{\underline{r}}_0$, X_m^0 and V_m^0 . The integration output is obtained at discrete UTC time intervals, according to the specified step size. The numerical integration methods used in ORBIT are :

- (i) a 4th order Runge-Kutta (R-K) starting procedure (see Appendix D.1), followed by
- (ii) an 8th order Adams-Bashforth (A-B) predictor-corrector method (see Appendix D.2).

The step sizes of these two methods, H_1 and H_2 , respectively, are supplied by the user. To start the predictor-corrector computation, the R-K formulae have to be used to provide 8 steps, in addition to the starting elements, at intervals of H_2 . This is equivalent to a total number of R-K steps, of

$$N_1 = \frac{H_2}{H_1} \times 8 \quad (3.7)$$

From the above equation, it is clear that the ratio H_2/H_1 represents the number of R-K steps per single A-B step, and this number must be an exact integer, since H_1 must be smaller than H_2 (see Section 2.4.2).

The step sizes must be chosen so as to obtain all the values (integrals) to the required accuracy. There is a stricter accuracy requirement for the satellite ephemeris than for the partials, therefore if the satellite positions are satisfactory, there is no need to check on the accuracy of V_m and X_m . A pair of appropriate step sizes have been established by trial and error

for LAGEOS, as

$$\begin{aligned} H_1 &= 15 \text{ seconds} \\ H_2 &= 120 \text{ seconds} \end{aligned} \quad (3.8)$$

On halving these step sizes and repeating the numerical integration, the satellite IF coordinates changed by less than 1 cm after a four-day arc. The advantage of the Adams-Bashforth method is now obvious. Here, one step of 120 seconds requires only two function evaluations (see Section 2.4.2), while in the R-K method the equivalent 8 steps of 15 seconds (=120 seconds) require 32 (8 x 4) function evaluations. So the predictor-corrector method is computationally up to 16 times cheaper than the R-K method. At the step sizes of (3.8), the number of R-K steps, N_1 , is 64 (from equation (3.7)), after which the predictor-corrector takes over.

In ORBIT, a check on the truncation error (ϵ_i) of the Adams-Bashforth routine is maintained at each integration step. Application of equation (2.144) for $n=8$, gives,

$$\epsilon_i = \frac{c_9}{b_9 - c_9} (r_i^{**} - r_i^*) \quad (3.9)$$

where r_i^*, r_i^{**} : predicted and corrected component of IF satellite position vector, \underline{r} , at i^{th} integration step.

This gives

$$\epsilon_i = 0.027 (r_i^{**} - r_i^*) \quad (3.10)$$

ORBIT evaluates ϵ_i for all the components of the satellite position vector, and outputs a warning message when its value exceeds $1 \mu\text{m}$.

The force model used in ORBIT is specially suited for LAGEOS (see Section 3.1). It consists of the following constituents (see Section 2.3), which are in accordance with the standards for the MERIT Campaign (Melbourne, 1983):

- (i) GEM 10 or GEM-L2 gravity model, truncated to degree and order 20
- (ii) third body attractions of moon, sun, Venus, Mars, Jupiter and Saturn
- (iii) lunar and solar solid earth tides (Wahr model)
- (iv) solar radiation pressure
- (v) an empirical along track acceleration
- (vi) ocean tides (Schwiderski model)

The approximate magnitudes of each of these constituents on LAGEOS, and their overall effects after 4 days, are tabulated in fig 3.III.

LAGEOS passes through the earth's shadow, and there is a facility in ORBIT to perform a shadow test (see Section 2.3.8) and switch off the solar radiation acceleration. If this facility is used it introduces sudden changes in the force model, which result in unacceptably high truncation errors (ϵ_1) in the Adams-Bashforth method. There are two possible solutions to this problem: in the first place, the solar radiation could be altered gradually. This would also represent more accurately the physical reality of the situation, as the satellite goes into the earth's penumbra. Secondly, the Adams-Bashforth corrector formula could be applied more than once. In this case the error expression in equation (3.10) would not apply, but it could still be used in order to ensure that successive 'corrected' values converge to a specified limit. The above two suggestions should be the subject of further study

Force Model Component	LAGEOS Acceleration (m/s ²)	Effect After 4 Days (m)
Earth Attraction	2.64	1.6×10^{11}
Moon Attraction	2×10^{-6}	1.2×10^5
Sun Attraction	9×10^{-7}	5×10^4
Lunar Earth Tides	2×10^{-8}	1.2×10^3
Solar Earth Tides	1×10^{-8}	600
Solar Radiation ($C_R = 1.2$)	3×10^{-9}	200
Along-Track Acceleration	3×10^{-12}	0.2
Planetary Attractions:		
(a) Jupiter	3×10^{-12}	0.2
(b) Venus	6×10^{-13}	0.04
(c) Saturn	2×10^{-13}	0.01
(d) Mars	3×10^{-14}	0.002
Ocean Tides	2×10^{-12}	0.1

Fig 3.III Typical Magnitudes of LAGEOS Force Model Constituents

(see Chapter 7). In the present situation the solar radiation cut-off is not used and, by making C_R an adjusted parameter, it is hoped to average out any errors arising from this.

The force model is also used to compute the derivatives of satellite acceleration, A_m (see Section 2.5.3 and Appendix C.2), with respect to the satellite state vector and force model parameters, from the position and velocity partials, X_m and V_m (equation (2.188)). This requires the evaluation of matrices $D1$, $D2$ and A_f . Since there is no need to account for air drag at the LAGEOS altitude, the only force model constituent which contributes to $D2$ is the along-track acceleration. This is very small (see fig 3.III) and it has been decided to ignore its effect on the partials and set the elements of $D2$ to zero. Equation (2.188) then becomes

$$A_m = D1 \cdot X_m + A_f \quad (3.11)$$

In Section 2.5.3 it was argued that the less significant components of the force model can be ignored in the evaluation of $D1$, without affecting the rate of convergence of the least squares solutions.

The components contributing to $D1$ in ORBIT, are:

- (i) all the terms of the geopotential,
- (ii) the third body attractions of the moon, sun, Venus, Mars, Jupiter and Saturn, and
- (iii) ocean tides.

The effect of ocean tides is insignificant, but these are included in the geopotential as corrections to the spherical harmonic coefficients (see Section 2.3.5), prior to the computation of $D1$.

In the evaluation of matrix A_f , the first six columns, which correspond to the elements of the state vector, are all zero. Each subsequent column corresponds to a force model parameter, p_1 (see Appendix C.2), and the total number of these, np , is specified by the user. There are options for computing partials for C_R , C_a , GM and for up to 12 normalized spherical harmonic coefficients, \bar{C}_n^m and \bar{S}_n^m .

3.3.2 Input Requirements

ORBIT requires input from up to 7 separate files, which are illustrated in fig. 3.IV. These contain the following information:

- (i) the input mode of the starting elements (INMOD). There are three available input modes:

- 1 : satellite state vector in IF coordinates. X_m^0 and V_m^0 given by (2.189) and (2.190).
- 2 : satellite state vector in EF coordinates. X_m^0 and V_m^0 as above.
- 3 : this mode operates as a continuation of a previous run of ORBIT. The last nine steps of the IF satellite ephemeris and of the matrices X_m and V_m from the previous run are input, and the R-K routine is by-passed. The partials in matrices X_m and V_m refer to the starting elements of the first run. This mode enables the user to perform several 'small' runs of the program, instead of one long run, and avoids the use of very large files.

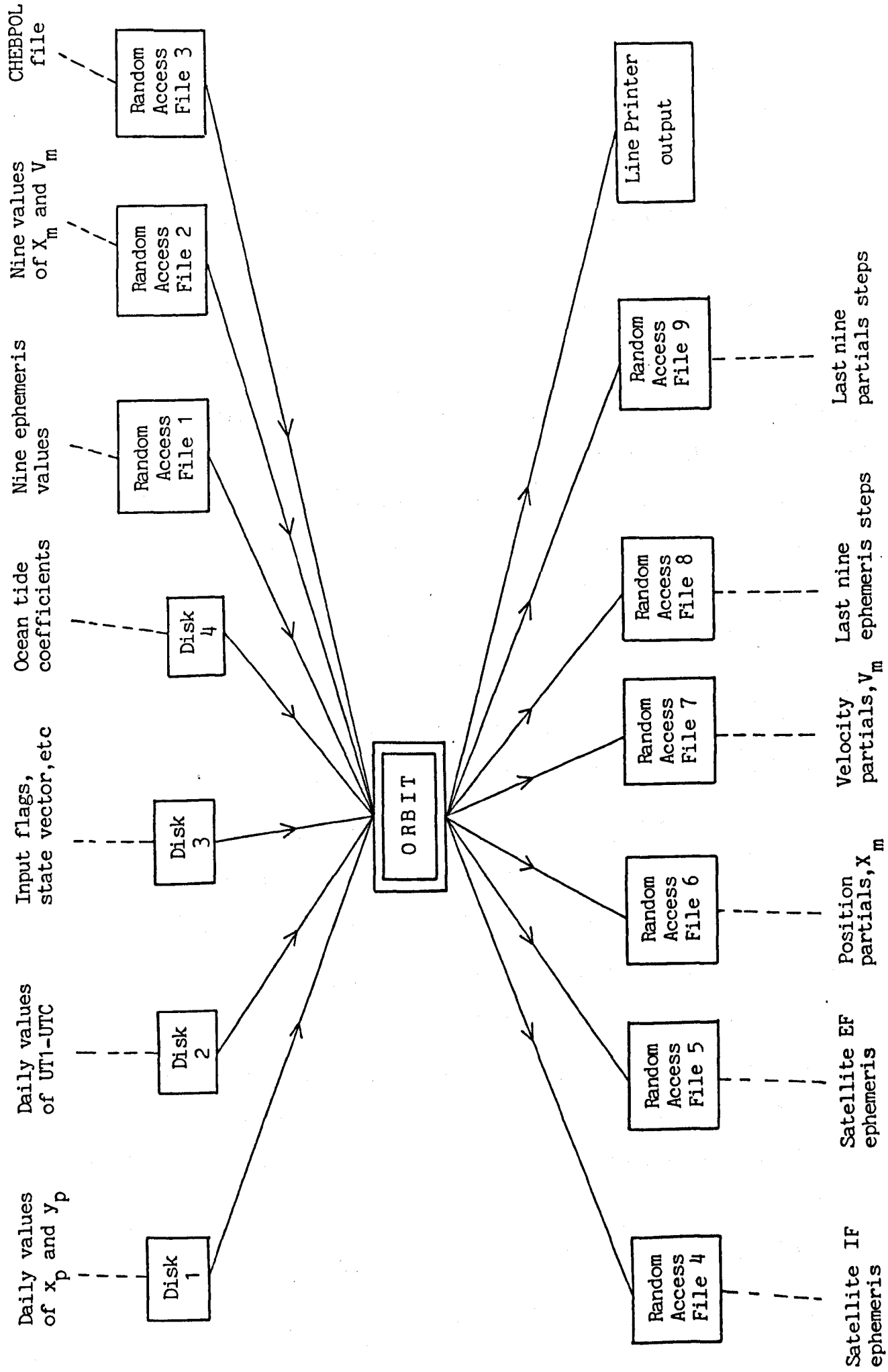


Fig 3.IV Data Input and Output for ORBIT

- (ii) the UTC epoch, t_0 , of the starting elements. The year, day number (of year), and time in hours, minutes and seconds is required.
- (iii) the satellite state vector in the appropriate reference frame (see (1)).
- (iv) the last nine integration steps of the previous run (only if INMOD=3).
- (v) the earth's angular velocity. This is required in order to convert the satellite IF velocity to the EF reference frame, relative to a set of rotating axes. Equation (2.95) gives the IF velocity relative to the rotating axes (\underline{v}_r), and this is rotated to the EF frame by (see section 2.2.4)

$$\underline{V}_r = P E N Q \underline{v}_r \quad (3.12)$$

\underline{v}_r is used in the computation of air drag (see Section 2.3.7), when an appropriate model is available.

- (vi) the satellite mass and cross-sectional area. These are only used in the computation of the solar radiation acceleration (see Section 2.3.8).
- (vii) the line printer output mode. This controls the amount of output supplied to the user.
- (viii) the integration step sizes, H_1 and H_2 (see Section 3.3.1), and the total number of steps, N_T , at intervals of H_2 . The number of R-K steps, N_1 , is given by (3.7). The number of A-B steps is

$$N_2 = N_T - 8 \quad (3.13)$$

- (ix) flags for the required force model partials (C_R , C_a , GM , \bar{C}_n^m , \bar{S}_n^m)
- (x) the degree and order of each of the spherical harmonic coefficients for which partials are needed
- (xi) the polynomial coefficients from the CHEBPOL file
(see Section 3.2)
- (xii) polar motion and each rotation flags. There are three options here :
 - 1 : x_p , y_p and/or UT1-UTC are linearly interpolated from the CHEBPOL file
 - 2 : ERP's are read from the CHEBPOL file, but they are treated as constant throughout the program run
 - 3 : daily values of x_p , y_p and/or UT1-UTC are input from separate files. This avoids the need to run CHEBPOL for a second time if the ERP's need to be updated
- (xiii) solar radiation cut-off flag. This operates the shadow test (see Section 2.3.8).
- (xiv) the values of C_R , C_a and GM
- (xv) the load numbers, k'_n , and the Schwiderski ocean tide coefficients for the tides described in Section 2.3.5

3.3.3 General Outline

This section summarizes the various tasks performed by ORBIT, and the reader is referred to fig 3.V which gives a simplified flow chart of the program.

After the various flags are input (see Section 3.3.2), the satellite starting elements are read according to the value of INMOD, and the numerical integration is started. There are two options at this stage :

(a) if the current run is a continuation of a previous program run, the R-K routine is by-passed and the predictor-corrector integration is started immediately. Nine consecutive values of \underline{r} , $\dot{\underline{r}}$, X_m and V_m , from the output of the previous run (see Section 3.3.4), are input. The partials in matrices X_m and V_m are given, in this option, with respect to the elements of the state vector which started the numerical integration of the very first program run. The input of the nine steps is from two random access files (see fig 3.IV). The first file consists of nine records, each containing the elements of the satellite IF position and velocity vectors for one integration step. The first number in each record is the UTC time, Δt_1 , elapsed since the epoch of the state vector. This is used to establish the UTC time at each subsequent integration step of the Adams-Bashforth routine (see fig 3.V). The epoch of the first step, for example, is given by

$$t = t_0 + 1 \times H_2 + \Delta t_9 \quad (3.14)$$

where t_0 : epoch of state vector

Δt_9 : time elapsed from t_0 to the epoch of the last ephemeris record of the previous run

H_2 : integration step size of A-B method

The second file also consists of nine records, each holding the elements of matrices X_m and V_m for one integration step.

(b) if this is the first run of ORBIT with a particular set of starting elements, the Runge-Kutta routine is initiated. This routine is allowed to progress for N_1 steps, computed using equation (3.7), at a step size of H_1 . It computes the equivalent of eight steps at a step size of H_2 , which are then used to start the Adams-Bashforth integration scheme.

The CHEBPOL data is initially read for the date of the satellite state vector. At each integration step a check is maintained on the date and when this changes, a new set of polynomial coefficients is retrieved. The coefficients enable ORBIT to compute the precession-nutation matrix, B , the nutation in longitude, $\Delta\phi$, the planetary and lunar coordinates, and the ERP's, from the known UTC time (see Section 3.2.2). Daily values of x_p , y_p and UT1-UTC can also be read from disk, without having to generate a new CHEBPOL file, if there is a need for them to be updated.

In the transformation between IF and EF reference frames, the rotation matrices P , E , N and Q are required (see Section 2.2). Since N and Q are combined into one matrix, B , the complete transformation from IF to EF coordinates (see Section 2.2.4), is

$$\underline{R} = D \underline{r} \quad (3.15)$$

where

$$D = PEB^T \quad (3.16)$$

The reverse transformation is

$$\underline{r} = D^T \underline{R} \quad (3.17)$$

ORBIT computes matrices D and D^T at every step, because they are both required when deriving the satellite acceleration and the partial derivatives which depend on the geopotential (see Section 2.3.2 and Appendix C). The normalized spherical harmonic coefficients, \bar{C}_n^m and \bar{S}_n^m , are held in a BLOCK DATA segment. There are currently two identical versions of ORBIT, one with the GEM10 and the other with the GEM-L2 (see Section 2.3.2) coefficients.

At every integration step the IF satellite acceleration, $\ddot{\underline{r}}$, and acceleration partials, A_m (see Section 2.5.3), are computed as follows:

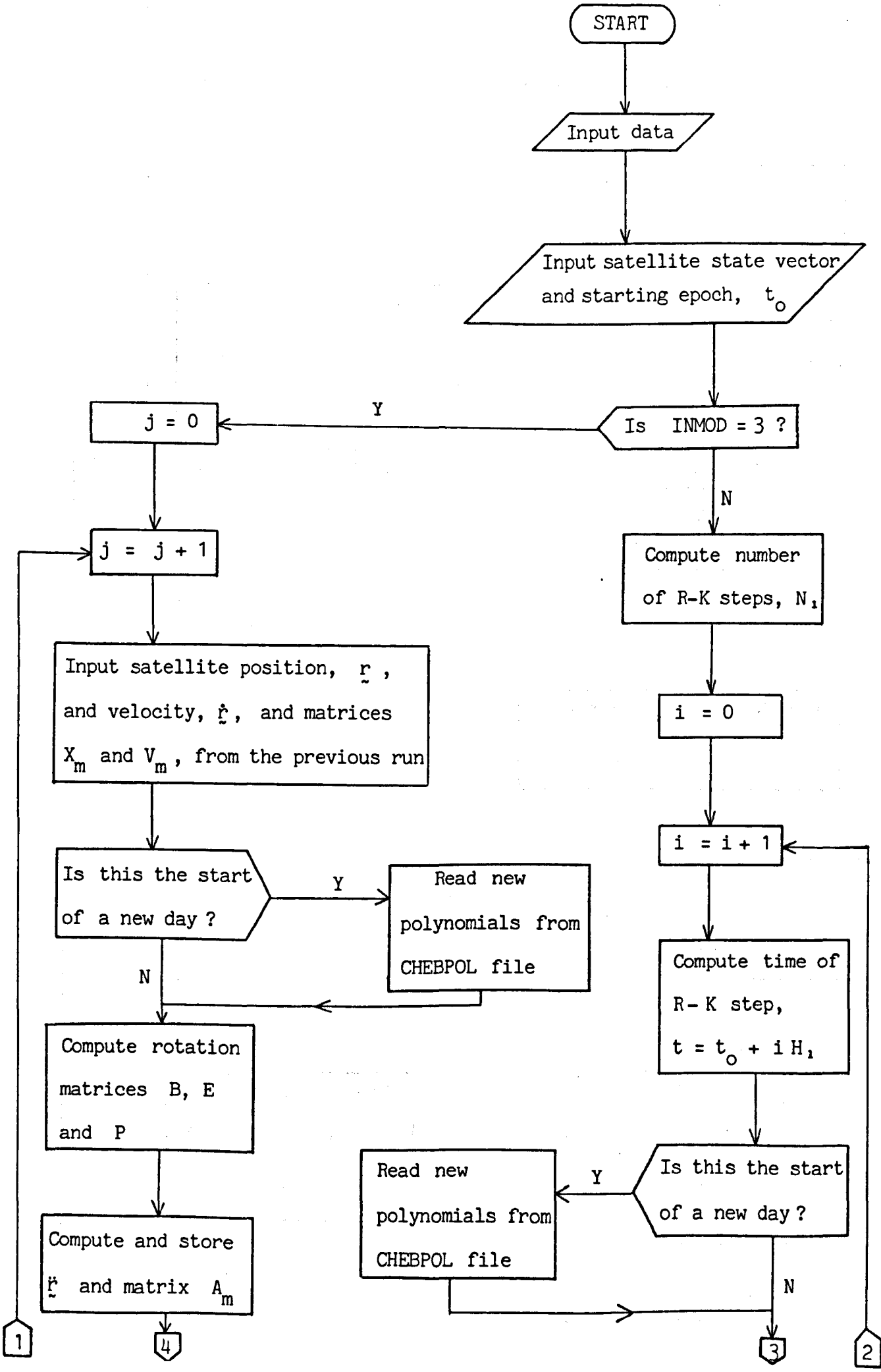
- (i) the normalized geopotential coefficients are corrected for ocean tides and for frequency dependent Love numbers (see Sections 2.3.4 and 2.3.5).
- (ii) the IF planetary and lunar coordinates are evaluated from the Chebyshev polynomials. These are used, along with the IF satellite position vector, \underline{r} , to compute the satellite acceleration arising from third body gravitational attraction (Section 2.3.3), solar radiation pressure (Section 2.3.8) and solid earth tides (Section 2.3.4). The elements of matrix $D1$ (see Appendix C.2) which depend on third body gravitational effects, are also computed. If required, the

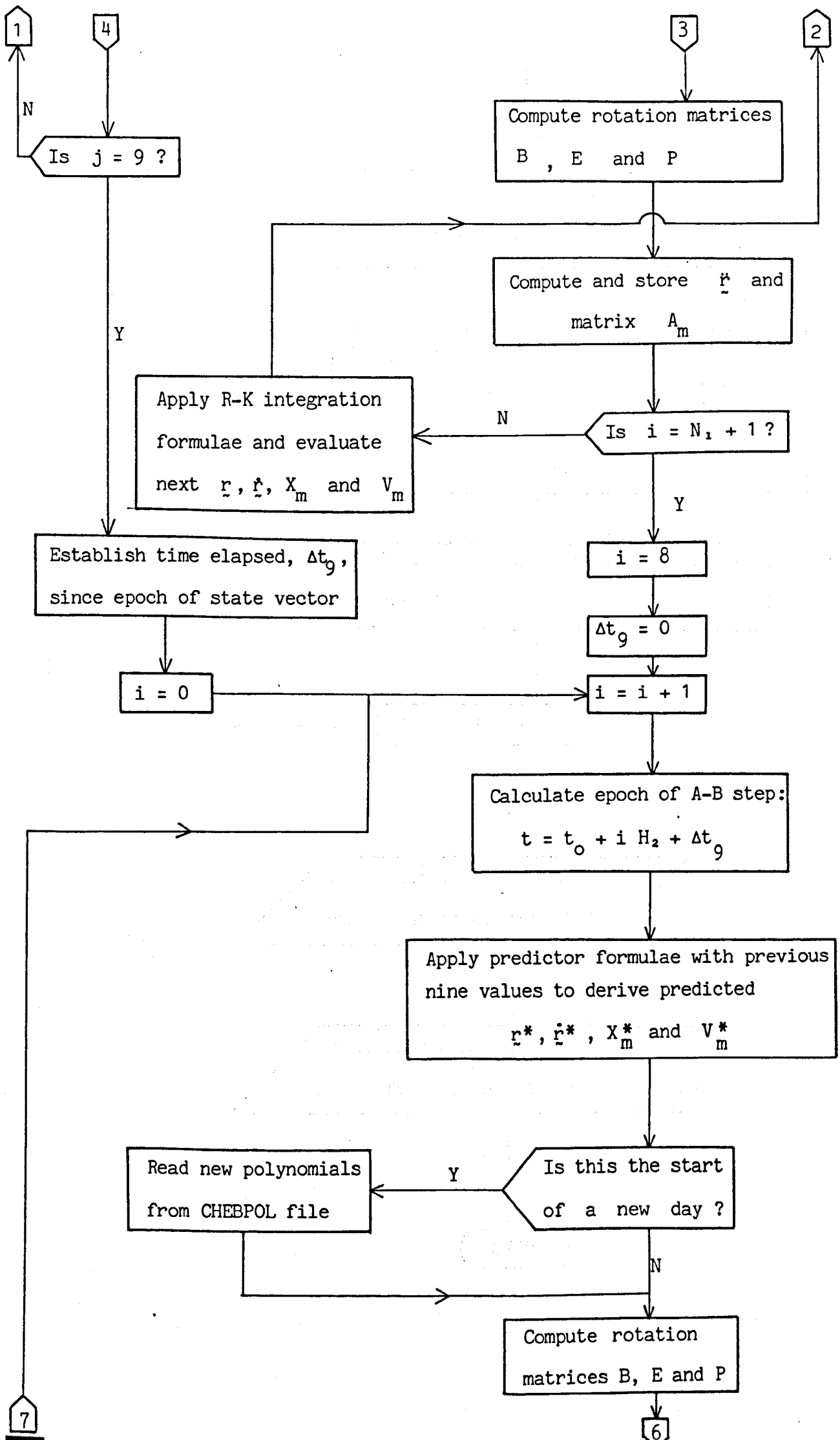
derivatives $\frac{\partial \ddot{\mathbf{r}}^*}{\partial \mathbf{C}_R}$ in matrix A_f are evaluated.

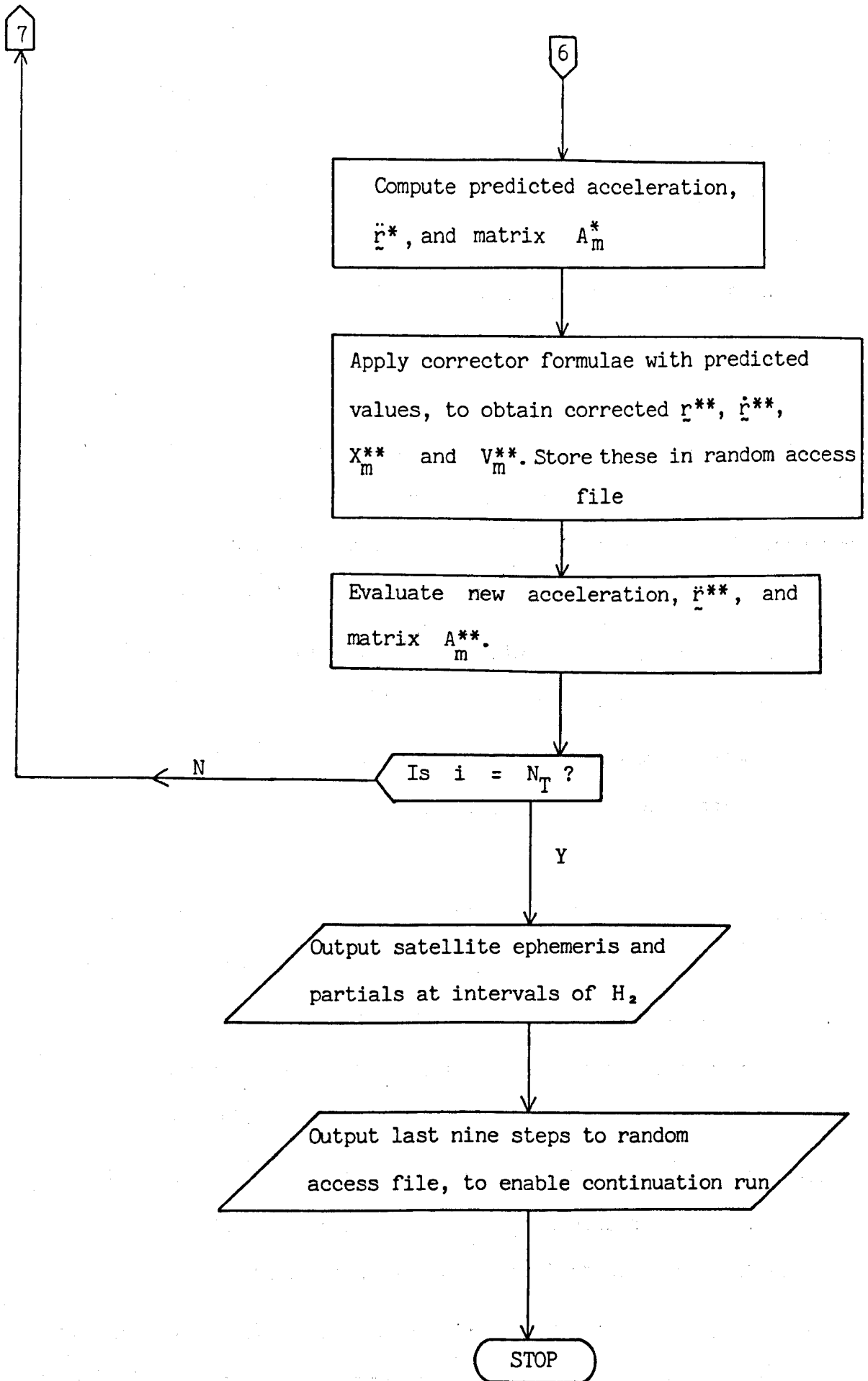
- (iii) the IF satellite coordinates are transformed to the EF reference frame, using equation (3.15), and from a cartesian to a polar representation (see Appendix B.2)
- (iv) the polar satellite coordinates and the geopotential coefficients are used to compute the EF acceleration, $\ddot{\mathbf{r}}$, and derivatives, $D1E$, arising from the earth's attraction (see Appendix C). These are rotated to the IF by making use of matrices D and D^T .
- (v) a specified number of the geopotential partial derivatives: $\frac{\partial \ddot{\mathbf{r}}^*}{\partial GM}$, $\frac{\partial \ddot{\mathbf{r}}^*}{\partial \mathbf{C}_n^m}$ and $\frac{\partial \ddot{\mathbf{r}}^*}{\partial \mathbf{S}_n^m}$ (see Appendix C.2) are evaluated and rotated to the IF. These constitute consecutive columns of matrix A_f
- (vi) the IF satellite velocity vector is used to compute the empirical along-track acceleration (Section 2.3.6) and partials $\frac{\partial \ddot{\mathbf{r}}^*}{\partial \mathbf{C}_a}$
- (vii) the program now calculates the resultant satellite acceleration, $\ddot{\mathbf{r}}$, and matrix $D1$, from all the appropriate force model constituents
- (viii) the acceleration partials, A_m , are derived from the position partials, X_m , and from $D1$ and A_f , using equation (3.11)

The program execution stops when the required number of integration steps have been computed. Details of the final output of ORBIT are given in the next section.

Fig 3.V ORBIT Flow Chart







3.3.4 Program Output

ORBIT produces a variable amount of line printer output and six random access output files (see fig 3.IV). These satisfy the present requirements of the Satellite Orbit Analysis Program (SOAP) for range data (see Section 3.4), as well as those of any future modifications of SOAP to deal with other types of data (see Section 2.5.1). The six files are:

- (i) the IF ephemeris file. This contains the IF satellite position and velocity vectors at UTC time intervals of H_2 . Each record is tagged with the time elapsed since the epoch of the starting elements.
- (ii) the EF ephemeris file. This is the same as (i) but the ephemeris is in the EF reference frame.
- (iii) the file of position partials. Each record contains successive values of matrix X_m , at the same time intervals as (i)
- (iv) the file of velocity partials. This contains successive values of matrix V_m .
- (v) the IF ephemeris file containing the last nine integration steps. This constitutes the last nine records of (i).
- (vi) the file of partials containing the last nine integration steps. Each record contains consecutive values of matrices X_m and V_m from the last nine integration steps.

Files (v) and (vi) are the ones used as input in a continuation run of ORBIT, based on the previous set of starting elements (see Section 3.3.2). The current version of SOAP makes use of files (i)

and (iii). Range is not a function of satellite velocity, so the velocity partials, V_m , (file (iv)), are not used. For the same reason, the satellite velocities in file (i) are ignored. The EF ephemeris file (file (ii)) would be essential in any future development of programs which only solve for station coordinates. Here the adjustment is performed entirely in the EF frame, avoiding the need to incorporate the rather complex precession and nutation theories.

3.3.5 Software Debugging

In any software development, the programs have to be exhaustively tested and debugged. In ORBIT, tests were conducted in order to answer the following three questions:

- (a) What is the effect of using Chebyshev polynomials?
- (b) What is the precision of the partials, X_m ?
- (c) How accurate is the ephemeris?

ORBIT was initially developed using an exact evaluation of matrix B, without making use of Chebyshev polynomials. This enabled the testing of two versions of the program, after the Chebyshev representation had been implemented. Comparison between the ephemeris output of the two programs showed no appreciable differences.

To test the precision of the partials, X_m , they were approximated by

$$\frac{\partial r_j}{\partial r_i^0} = \frac{\Delta r_j}{\Delta r_i^0} \quad (3.18)$$

where Δr_i^0 : small shift in an element of the satellite state vector ($\underline{r}_0, \dot{\underline{r}}_0$)

Δr_j : change in one of the components of the satellite position vector, caused by Δr_i^0

ORBIT was first executed using a typical set of LAGEOS starting elements. Six subsequent runs were performed, introducing shifts in each of the components of the satellite state vector in turn, of 1 m for \underline{r}_0 and of 1 mm/s for $\dot{\underline{r}}_0$. The numerical derivatives of (3.18) were computed, and the agreement between these and the actual partials, X_m , was good to 3 significant figures.

The precision of the numerical integration was tested by halving the integration step sizes and repeating the run (see Section 3.3.1). Unfortunately there can be no internal checks to detect gross errors in the formulation of the force model and the computation of the ephemeris. However, an excellent check was provided by the comparison of the ephemeris output of ORBIT with that of an orbit integration program developed by Dr. A. T. Sinclair of the Royal Greenwich Observatory. This uses a similar force model, but different numerical integration procedures. A common set of LAGEOS starting elements was used to independently generate a satellite ephemeris spanning four days. The maximum difference in any of the components of satellite position between the two programs was less than 1.5 cm. A final check was provided, with the use of ORBIT in an adjustment with SLR observations of LAGEOS. The solutions for tracking station coordinates were in good agreement with those obtained by the NASA Goddard Space Flight Center (see Section 4.2.3).

3.4 THE SATELLITE ORBIT ANALYSIS PROGRAM (SOAP)

3.4.1 Program Input and Output

SOAP uses the SLR tracking observations and the output of ORBIT and CHEBPOL in order to perform the least squares adjustment. It treats the observations as two-way ranges (see Section 2.5.1) and can work either with the full data set or with normal points from DATPAK (see Section 3.1). The input-output configuration is illustrated in fig 3.VI. The main input requirements of SOAP are :

(i) the mode of operation of the program. There are three available modes :

- 1 : the program reads the observations and performs the least squares adjustment,
- 2 : the program forms the observation equations and stores them in an unformatted file,
- 3 : the observation equations from a number of runs under mode 2 are retrieved and the program performs the least squares adjustment.

Modes 2 and 3 are necessary for large data sets where, to avoid dealing with very big files, the adjustment is performed in stages. The data is first divided into smaller (say 4-5 day) batches. ORBIT is then executed by initiating successive continuation runs from one set of starting elements (see Section 3.3.3), each run covering the duration of one batch of observations. The observation equations for each batch are formed and stored (mode 2) and, when the whole data set has been processed, they are retrieved and adjusted for the various unknowns (mode 3).

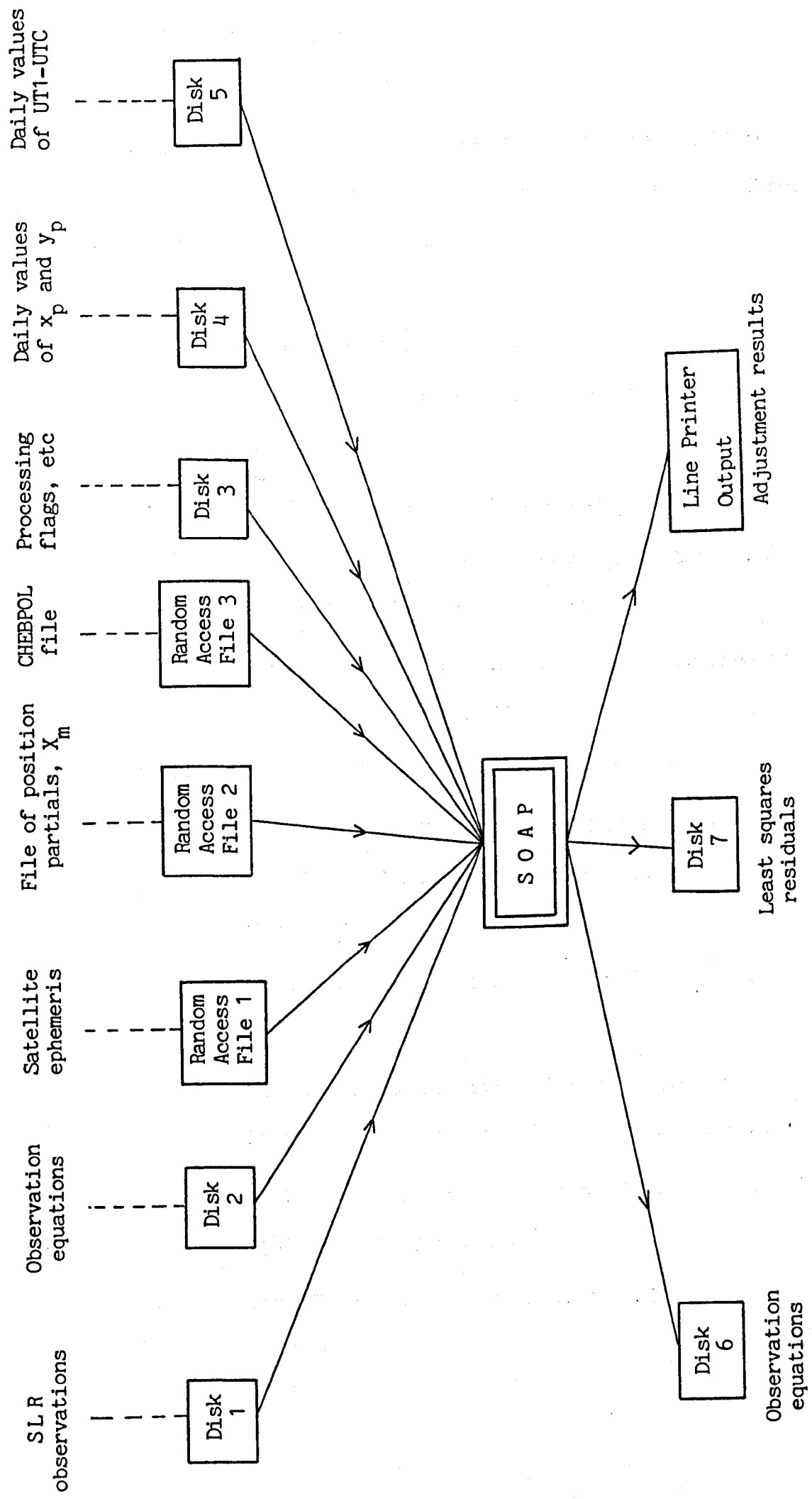


Fig 3.VI SOAP Input and Output

- (ii) the output mode. This controls the amount of line printer output.
- (iii) the centre of mass correction to the satellite. This corrects the observed ranges to the centre of mass of the satellite.
- (iv) the semi-major axis and the reciprocal of the flattening ($\frac{1}{f}$) of the spheroid in which the tracking station coordinates are given (see Appendix B.3)
- (v) the number of tracking stations
- (vi) the approximate coordinates of each station, in longitude, latitude and height
- (vii) the epoch of the satellite state vector used in ORBIT (see Section 3.3)
- (viii) the IF elements of the state vector
- (ix) the step size, H_2 , of the predictor-corrector used in ORBIT. The ephemeris and partials are given at intervals of H_2 .
- (x) the number and identity of each station to be held fixed
- (xi) the number and identity of each station whose longitude is to be held fixed (see Section 2.5.4)
- (xii) a flag which allows the program to fix the satellite orbit (starting elements)
- (xiii) the IF satellite ephemeris file (from ORBIT)

- (xiv) the file of IF satellite position partial derivatives, X_m , with respect to the starting elements and to the required force model parameters (from ORBIT)
- (xv) the file of polynomial coefficients from CHEBPOL
- (xvi) the number of force model unknowns and their provisional values. There are options for solving for GM, C_R , C_a and for a number of normalized geopotential coefficients, \bar{C}_n^m and \bar{S}_n^m (see Section 3.3.1). The program is only able to solve for parameters for which partials have been evaluated by ORBIT.
- (xvii) the number of ERP unknowns. The program can solve for x_p and y_p and/or UT1-UTC.
- (xviii) the input mode of the ERP's. There are three options similar to those used in ORBIT (see Section 3.3.2). These are :
 - 1 : x_p , y_p and/or UT1-UTC are linearly interpolated to the epoch of observation from the values in the CHEBPOL file
 - 2 : ERP's are initially read from the CHEBPOL file and kept constant throughout the adjustment. This is the mode that is used when solving for the ERP's
 - 3 : daily values of x_p , y_p and/or UT1-UTC are input from separate files
- (xix) the SLR tracking observations (if the program is operating in modes 1 or 2)

- (xx) the file of observation equations (if operating in mode 3)

When SOAP has been executed in modes 1 or 3, the following output is available :

- (i) a line printer output, which includes the values of all the adjusted parameters with their a posteriori standard errors. This can be varied in accordance with the output mode to include the normal equations and the full covariance matrix (see Section 2.5.2)
- (ii) a file of post-adjustment observation residuals, which can be plotted using a specially developed graph-plotting program

When SOAP has been executed in mode 2 the least squares adjustment is not performed and the only output is an unformatted file containing the stored observation equations (see fig 3.VI).

3.4.2 General Description

SOAP performs the least squares adjustment according to the principles outlined in Section 2.5. The SLR tracking observations have already been corrected for tropospheric effects, using the Marini and Murray model (Marini and Murray, 1973), by the individual tracking facilities. This eliminates the need to include a tropospheric model in SOAP, unless a different model is to be adopted.

After reading the various input flags, if SOAP is operating in modes 1 or 2 it enters a loop where the observations are input

in chronological order. These provide the following information :

- (i) the identify of the tracking station
- (ii) the UTC time of firing of the laser pulse
- (iii) the observed two-way range

The Chebyshev polynomial coefficients are initially read from the CHEBPOL file for the date of the satellite starting elements.

For each observation, if the date changes, a new set of coefficients is read (see also Section 3.3.3). These enable SOAP to compute the coordinates of the moon and sun, in order to correct the tracking station coordinates for the effects of earth tides (see Section 2.3.4).

The next task of the program is to iteratively establish the time of observation, which is the epoch that the laser pulse reaches the satellite, and the computed value of the two-way range. First, the time of observation is assumed to equal the time of firing, t_F . The IF satellite position, from the ORBIT ephemeris file, is interpolated to this epoch by means of Everett's formulae (see Appendix E.2) up to the fourth central difference. The range, l_1 , between satellite and station is then computed using equation (2.166). This requires that the station coordinates be rotated to the IF by making use of matrix D (see equation (3.16)), for the epoch of firing. The range, l_1 , gives a second approximation of the epoch of observation, by

$$t_2 = t_F + \frac{l_1}{c} \quad (3.19)$$

where c : speed of light in vacuo.

The new time, t_2 , is used to derive a new IF satellite position vector and a new range, l_2 , which in turn gives

$$t_3 = t_F + \frac{l_2}{c} \quad (3.20)$$

The process is repeated until successive values of t_i agree to a specified limit, which is determined by the satellite speed. For example, since LAGEOS travels at 5700 m/s, the time must be known to within 200 nanoseconds for an ephemeris accuracy of 1mm. When the time of observation has been established, the process is reversed and an iterative procedure adopted in order to estimate the time of return, t_R , of the pulse to the tracking station. The difference between t_R and t_F , scaled by the speed of light, is the computed value of the two-way range.

SOAP now computes the observation equation coefficients (partial derivatives) of equation (2.168). These are evaluated for the various unknowns according to the principles outlined on Section 2.5.3. In SLR the measurement is a two-way range and, to account for this, the observation equation coefficients are multiplied by 2. The coefficients for the satellite state vector and for the required force model unknowns are derived from the partials in matrix X_m , which are interpolated to the observation epoch (from the ORBIT file) using Everett's formulae up to the second central difference.

At this stage, the program can follow one of three routes :

- (i) if it is operating in mode 1, then the contribution of each observation is added to the normal equations and the observation equation placed in a temporary store
- (ii) in mode 2, the observation equations are stored in an unformatted file and the program execution stops
- (iii) in mode 3, the observation equations are input from the unformatted file, and the normal equations are formed.

The observation equations are not weighted, thus implying an a priori standard error of 1 m for each normal point.

When all the satellite tracking observations have been processed, and if the program is in modes 1 or 3, artificial observation equations similar to (2.169) are introduced in order to hold certain quantities fixed. Options exist to fix the satellite orbit (starting elements) and the coordinates of any of the tracking stations. A selected number of station longitudes (see Section 2.5.4) can also be held fixed.

The normal equations are solved using Choleski's method of triangular decomposition and the solutions are output on a line printer. It is also possible to output the normal equations and the full covariance matrix. After the solution, the observation equations are recovered from the temporary store and the least squares residuals are evaluated. These are output in a file and can be used by a plotting program to produce graphs, in order to study any residual trends (see Section 4.2.2).

It is worth noting that in the OD observation equations (Section 2.5.3) the satellite unknowns are in the IF, while the tracking station unknowns are in the EF reference frame. This requires that, in order to form the observation equations, the satellite and station coordinates must be available in both reference frames. For this reason, at every observation epoch SOAP evaluates matrices D and D^T from the polar motion, P , earth rotation, E , and precession-nutation, B , matrices. The elements of B for the UTC observation epoch are derived from the Chebyshev coefficients, as described in Section 3.2.

CHAPTER FOUR

ANALYSIS OF SATELLITE TRACKING DATA AND APPLICATIONS OF ORBIT DETERMINATION

4.1 Introduction

LAGEOS tracking data are being provided for Nottingham University by the Royal Greenwich Observatory, which receives them from the NASA Goddard Space Flight Center (GSFC). For the purposes of the OD software development, two four-day sets of data were used in order to achieve the following aims:

- (i) to test the software and detect any errors,
- (ii) to establish the precision with which the unknown quantities can be determined,
- (iii) to establish the effects of the various force model parameters (see Section 4.2.2),
- (iv) to test the suitability of different geopotential models (see Section 4.2.2),
- (v) to recover tracking station coordinates and to compare the solutions with those of other computing centres (see Section 4.2.3),
- (vi) to test the ability of the OD process to recover the ERP's (see Sections 4.2.4 and 4.2.5).

The last point is important in view of the MERIT main campaign (Wilkins, 1980) and the hope that the Nottingham software (see Chapter 3) will be used to analyse and report on the data collected from it.

A discussion of the various applications of orbit determination is given in Section 4.3.

4.2 RESULTS OF ANALYSIS

4.2.1 LAGEOS Tracking Data

The results in this chapter are based on two sets of LAGEOS tracking data, observed between 2-5 September and 2-5 December, 1980. The contributing tracking stations and their geodetic coordinates, from the GSFC SL5 geodetic solution (Christodoulidis et al, 1982), are listed in fig 4.I. The locations of the stations are also shown on a map in fig 4.II. A station-by-station and day-by-day breakdown of the normal points (see Section 3.1) derived from the two data sets is shown in fig 4.III. The total number of 'raw' observations in the eight days was 41769. After DATPAK, these observations were reduced to 1001 normal points (see fig 4.III), each representing a one minute set of data. This amounts to a data compression of approximately 40:1 and illustrates the advantage of using normal points. The lasers used to take these observations were generally 2nd generation, with a firing rate of around 1 pulse per second and a measurement accuracy of 10-20 cm. A typical firing rate of the most advanced, 3rd generation laser tracking facilities is around 10 pulses per second, with a measurement accuracy of 3-5 cm (Sharman, 1982). The use of normal points with these most recent systems is even more important, since there is a ten-fold increase in the amount of data generated.

The observations were corrected for tropospheric delays according to the Marini and Murray model (Marini and Murray, 1973). This has an estimated accuracy of between 1-3 cm. The other correction applied to the data was a centre of mass correction. This is necessary since the measurement is made to the retroreflectors on the surface of the satellite, while the ephemeris refers to its centre

STATION LOCATION	CODE	LONGITUDE	LATITUDE	HEIGHT (m)
Greenbelt, Md., USA	STAL 7063	283° 10' 19" 800	39° 01' 13" 358	15.25
Haystack, Mass, USA	HAYS 7091	288° 30' 44" 345	42° 37' 21" 682	88.35
Owens Valley, Ca., USA	OWEN 7114	241° 42' 22" 221	37° 13' 57" 209	1174.59
Goldstone, Ca., USA	GOLD 7115	243° 12' 28" 954	35° 14' 53" 898	1035.16
Pasadena, Ca., USA	PASA 7896	241° 49' 39" 724	34° 12' 20" 023	437.89
Yarragadee, Australia	YARR 7090	115° 20' 48" 111	-29° 02' 47" 412	237.41
American Samoa	PAGO 7096	189° 16' 30" 356	-14° 20' 07" 519	45.14
Maui, Hawaii	HAWA 7120	203° 44' 38" 107	20° 42' 27" 391	3064.18

Spheroid : $\frac{1}{f} = 298.255$

$a = 6378144.11 \text{ m}$

Fig 4.I SL5 Coordinates of LAGEOS Tracking Stations

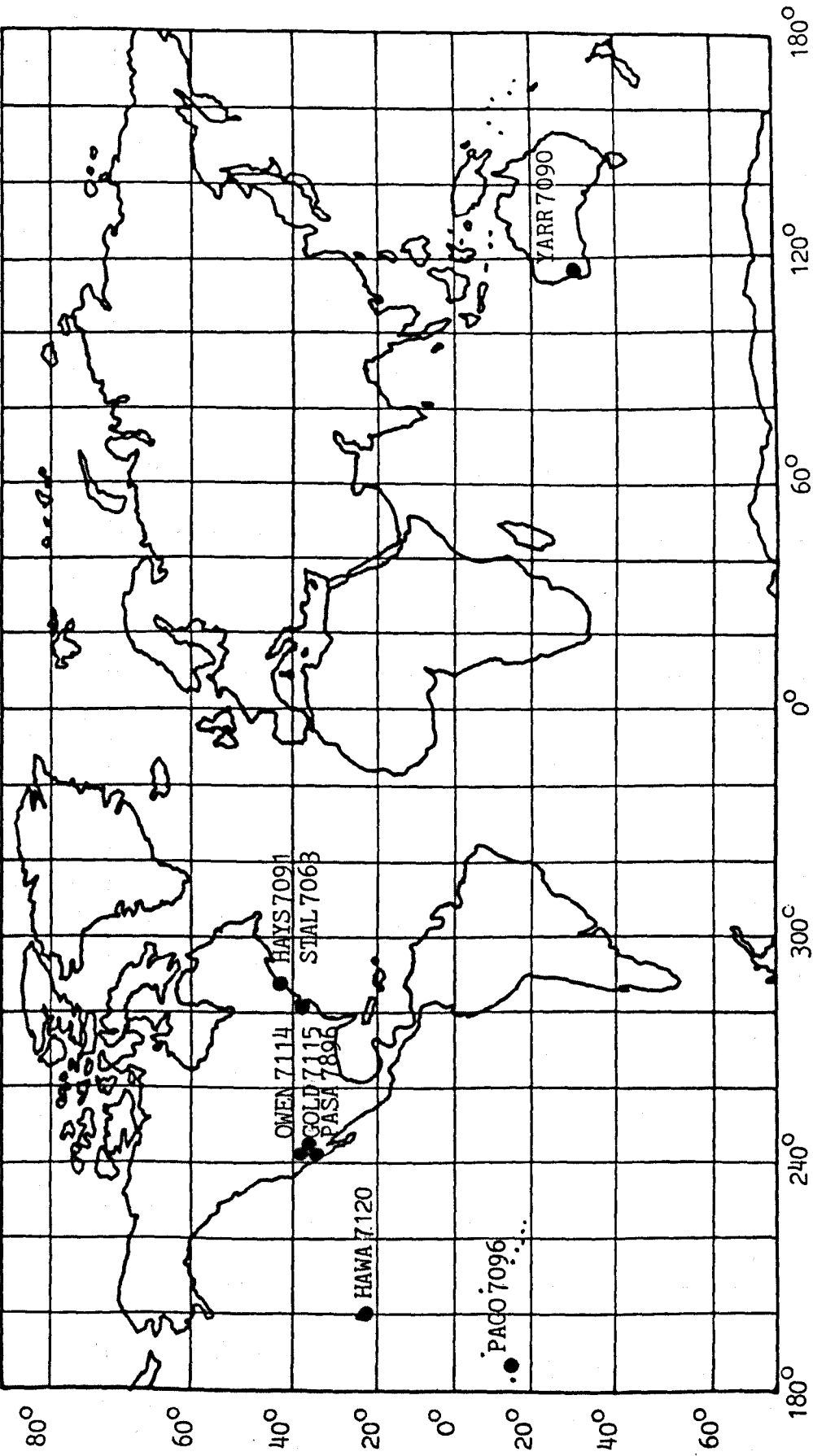


Fig 4.II LAGEOS Tracking Sites

STATION	DATA SET 1 (DS1)					DATA SET 2 (DS2)					
	NUMBER OF NORMAL POINTS					NUMBER OF NORMAL POINTS					
	2/9/80	3/9/80	4/9/80	5/9/80	TOTAL	2/12/80	3/12/80	4/12/80	5/12/80	TOTAL	
STAL 7063	35	9	35	-	79	21	4	2	-	27	
HAYS 7091	15	13	58	-	86	-	-	-	-	-	
OWEN 7114	-	-	-	-	-	-	-	-	28	28	
GOLD 7115	-	9	39	36	84	-	-	3	31	34	
PASA 7896	-	-	-	-	-	26	16	-	-	42	
YARR 7090	92	88	99	-	279	33	40	91	14	178	
PAGO 7096	-	43	-	-	43	-	-	-	-	-	
HAWA 7120	18	39	-	17	74	24	4	15	4	47	
Total for Data Set 1					645	Total for Data Set 2					356

Fig 4.III Distribution of Normal Points for the Two Data Sets

of mass. For LAGEOS (diameter 60 cm), the value of this correction was found, after calibration, to be 24 cm.

From the table in fig 4.III it can be seen that only a small part of the 4-day LAGEOS arcs is covered in each data set. In the most optimistic case, assuming that the observations were not simultaneous and bearing in mind that each normal point represents one minute of data, only 11% of the 4-day arc was covered by Data Set 1 (DS1) and only 6% by Data Set 2 (DS2). It is also seen that station YARR 7090 dominated both data sets, contributing 43% of the normal points in DS1 and 50% in DS2. It can be argued that since the coverage of the satellite orbit was not continuous, the data is not very suitable for a precise orbit determination. This would be especially true for a low altitude satellite which is significantly affected by air drag. However, since LAGEOS is in a high and stable orbit, it was considered that the data was sufficient to give an indication of the capabilities of orbit determination and to achieve the objectives listed in the previous section.

4.2.2 Effect of Force Model Parameters and Geopotential Models

The first adjustments of the LAGEOS data were performed with the version of ORBIT (see Section 3.3) which uses the GEM 10 (Lerch et al, 1979) geopotential model. The values of the most important parameters used in the orbit integration are listed below:

Geocentric gravitational constant (GM)	:	398600.5 km ³ s ⁻²
Earth's equatorial radius (a)	:	6378137 m
Solar radiation coefficient (C _R)	:	1.17
Along-track acceleration (C _a)	:	-2.9 x 10 ⁻¹² ms ⁻²

All the remaining force model constituents were in accordance with the MERIT standards (Melbourne, 1983) and they are summarized in Section 3.3.1.

Two sets of approximate LAGEOS IF satellite starting elements, for the two data sets, were provided by the Royal Greenwich Observatory. These elements were for 0.^h0 UTC of 2nd September and 2nd December, 1980. Using the starting elements, two 4-day LAGEOS arcs were generated by ORBIT with the integration step sizes of equations(3.8). Various solutions were performed by SOAP, and post-adjustment residuals were plotted. In the solutions described in this section the BIH values of x_p , y_p and UT1-UTC were linearly interpolated to the observation epochs. The only other constraint was a fixed longitude of station STAL7063 at the value given in fig 4.I.

In the first adjustment the unknowns were the EF coordinates of the six tracking stations and the IF elements of the satellite state vector. The least squares residuals were found to exhibit systematic trends, and a typical plot from DS1 is given in fig 4.IV. This graph shows that the residuals from the normal points of station STAL7063 exhibited a parabolic trend. It also shows that the normal points from HAYS7091 were rather noisy, with a scatter of around 30 cm. The root-mean-square (rms) residual of all the normal points, σ_R , was 27 cm.

To obtain the full precision of the orbit determination, it is necessary to isolate the causes of the systematic trends and find suitable models in order to remove them. The next adjustment solved for GM, in addition to the station coordinates and to the elements of the state vector. Fig 4.V shows that the parabolic pattern from STAL7063 was removed, but was replaced by a linear trend in the residuals. There was also a drop in σ_R , to 21 cm, implying an

improvement in the solution. It is obvious that the value of GM in the first adjustment was incompatible with the speed of light (299792500ms^{-1}) that was used to scale the ranges (see Section 2.5.4), resulting in the parabolic trend of fig 4.IV.

The causes of the linear trend in the residuals had now to be established. Two additional parameters, C_R and C_a , were allowed to adjust. Fig 4.VI shows that most systematic trends from STAL7063 were removed, and this is reflected by a drop in σ_R , to 17 cm.

The residuals from most other tracking stations in the same adjustment still exhibited some systematic patterns, as shown by the plots in figs 4.VII and 4.IX. At this stage, any residual trends could be explained by deficiencies in the geopotential model. A new satellite orbit was thus integrated, this time by replacing GEM10 with the specially derived LAGEOS model, GEM-L2 (Lerch et al, 1983). Another adjustment was performed, solving for the station coordinates, the satellite starting elements, GM , C_R and C_a , and the residuals were plotted. Figs 4.VIII and 4.X show the plots from the new adjustment for the same satellite passes as in figs 4.VII and 4.IX respectively. It is clear that most systematic trends have now been removed. This was also true of all other passes, and it can be concluded that GEM-L2 should be used in all adjustments with LAGEOS observations. The rms range residual of all the normal points, σ_R , was 10 cm for both DS1 and DS2 with GEM-L2, which is within the accuracy of the 2nd generation laser tracking facilities.

The final solutions for the various force model parameters, from the two data sets, are tabulated in fig 4.XI. The values of GM were highly repeatable and they depended on the adopted speed of light. Similar values have been reported independently by other

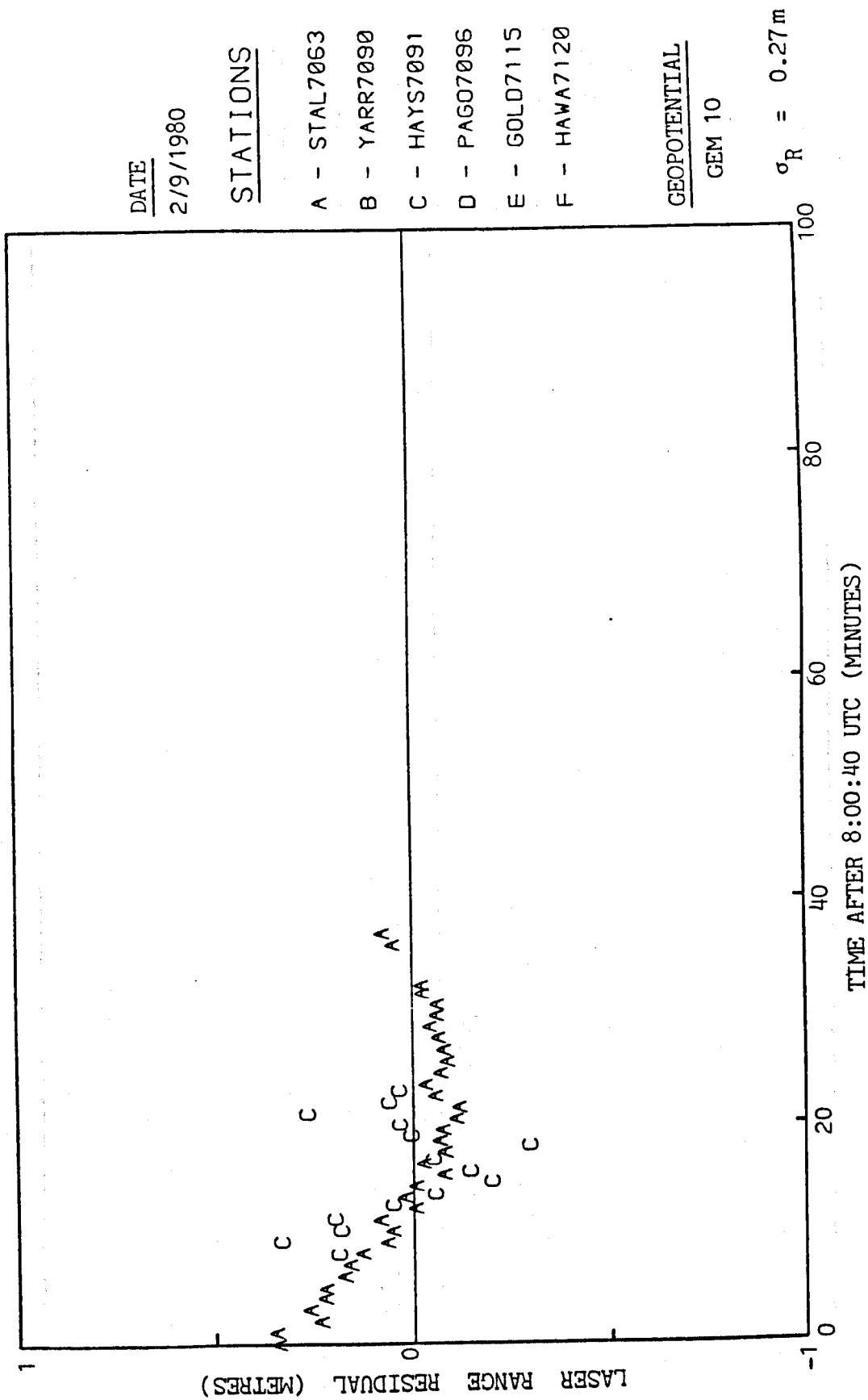


Fig 4.IV Range Residuals From Solution for Satellite State Vector and Station Coordinates

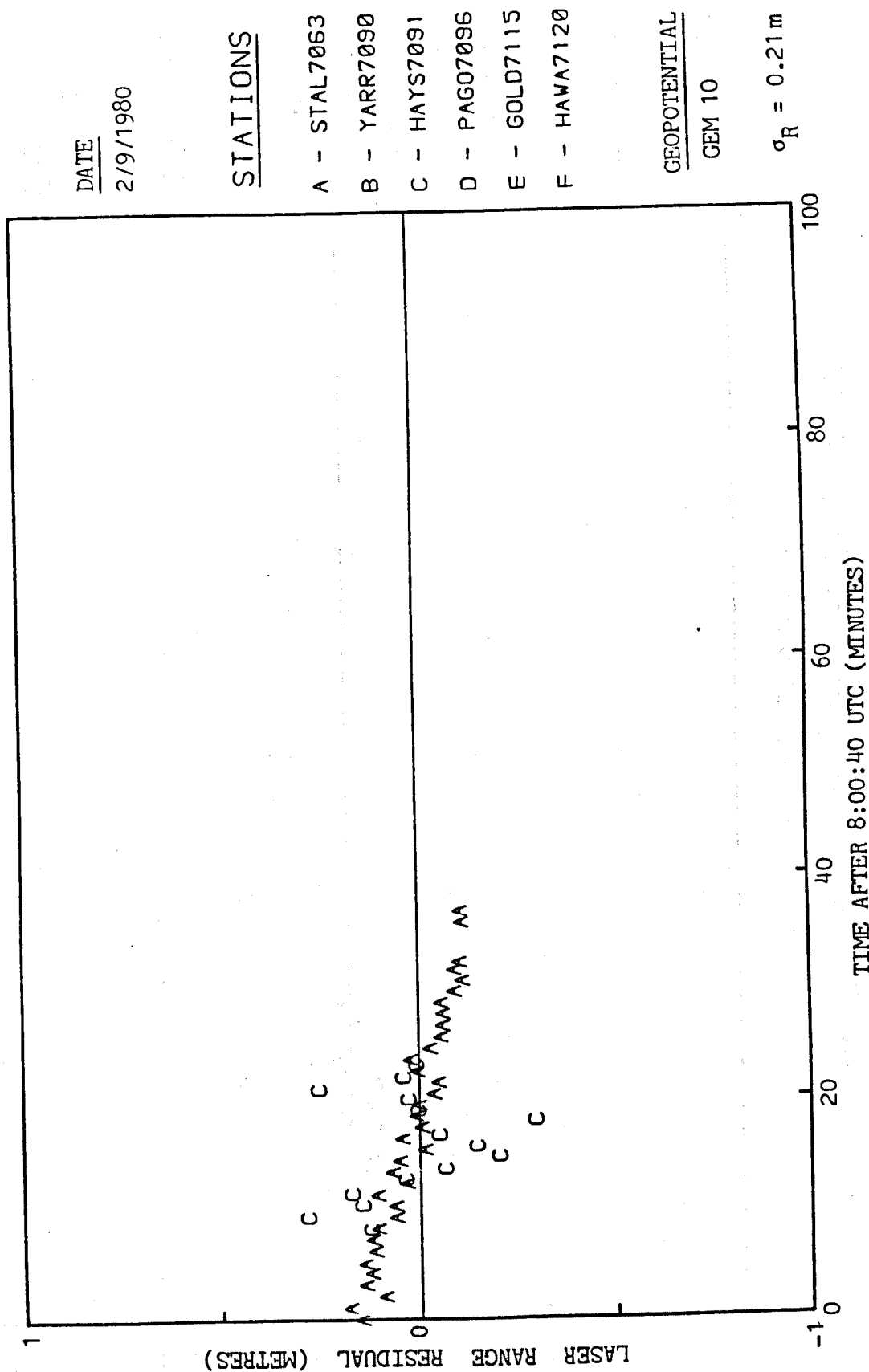


Fig 4.V Range Residuals From Solution for Satellite State Vector, Station Coordinates, and GM

DATE

2/9/1980

STATIONS

- A - STAL7063
- B - YARR7090
- C - HAYS7091
- D - PAG07096
- E - GOLD7115
- F - HAWA7120

GEOPOTENTIAL

GEM 10

$\sigma_R = 0.17 \text{ m}$

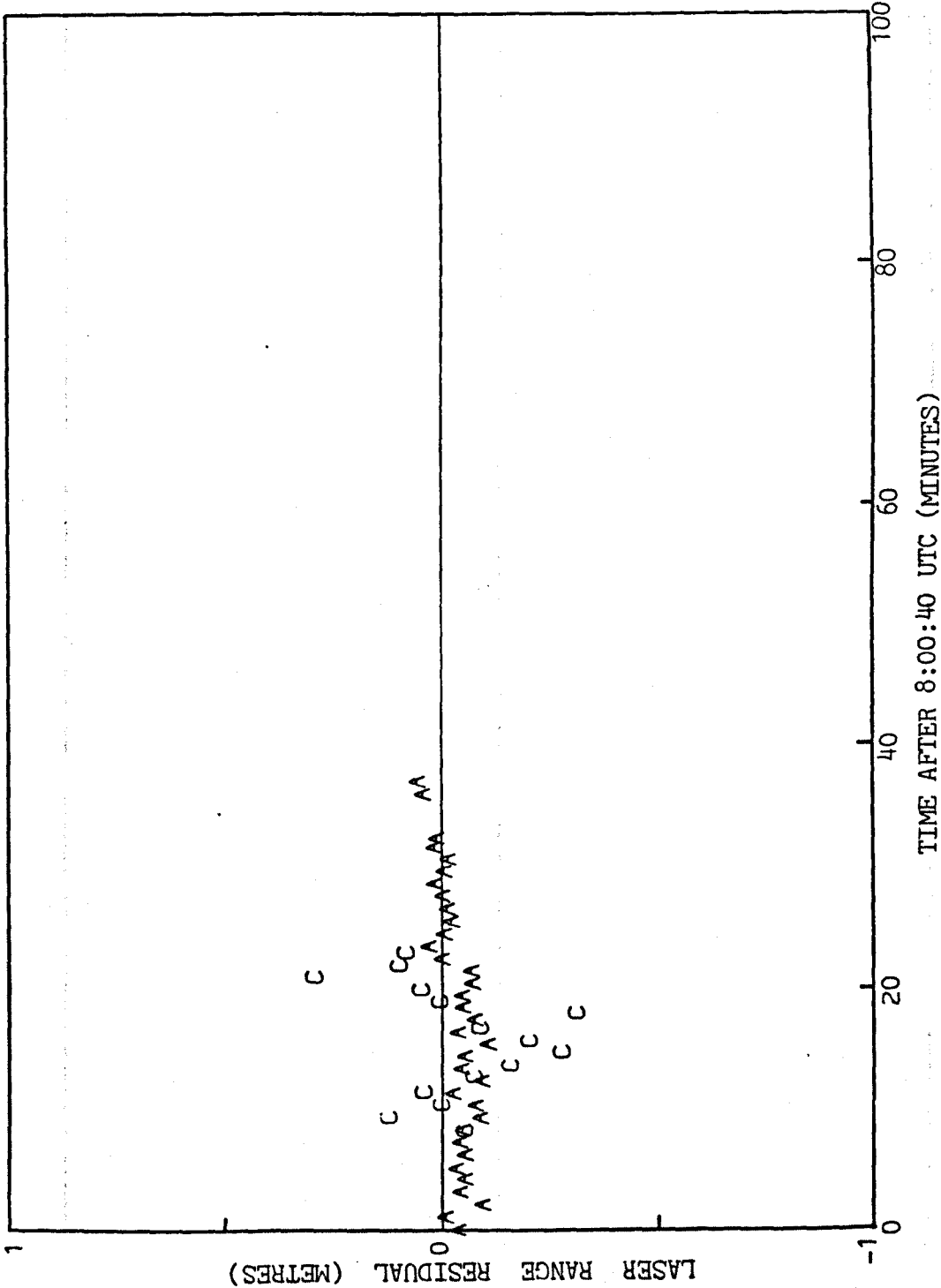


Fig.VI Range Residuals From Solution for Satellite State

Vector, Station Coordinates, GM, C_R , and C_a

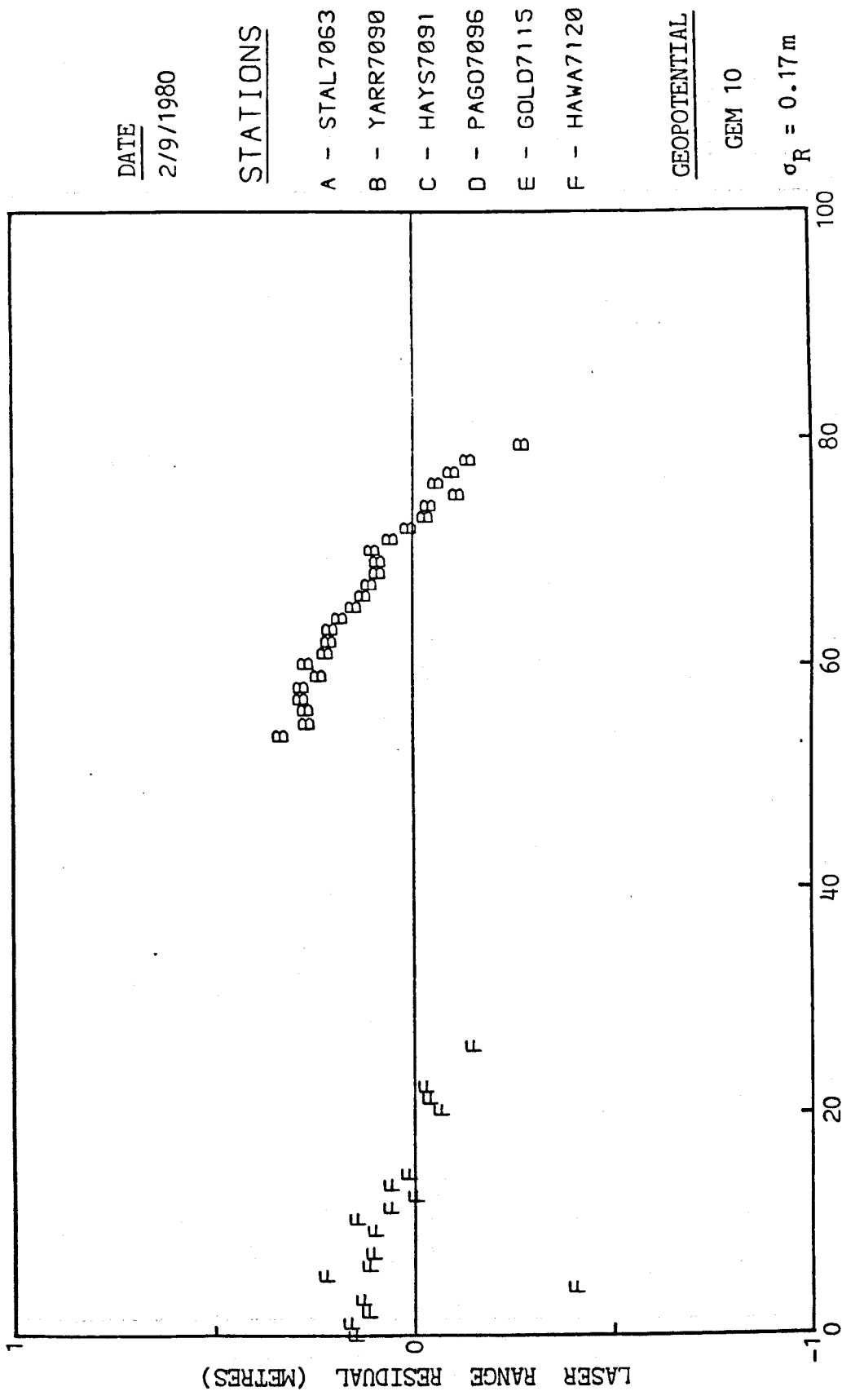


Fig 4.VII Range residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R, and C_a

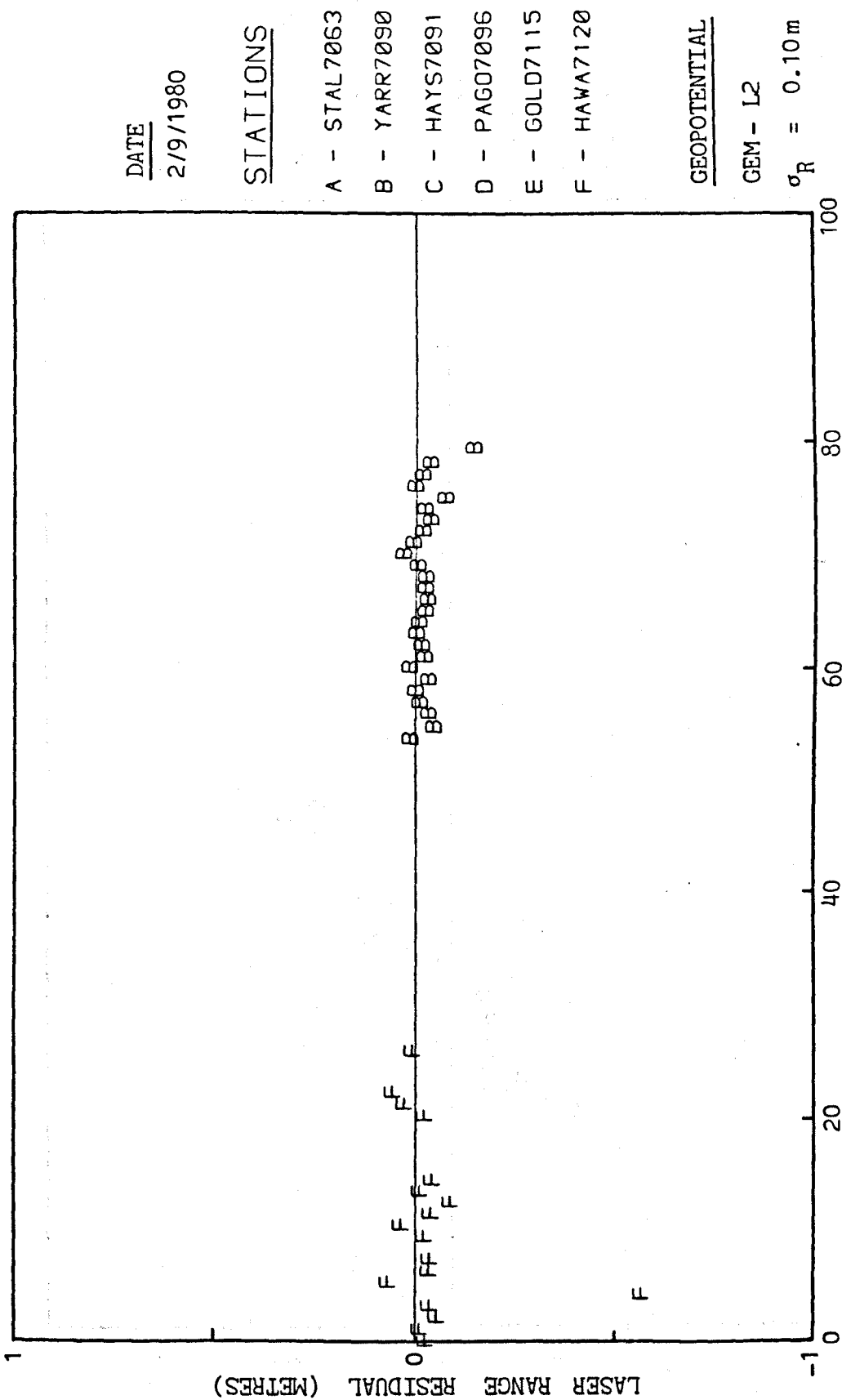


Fig 4.VIII Range Residuals From Solution for Satellite State
Vector, Station Coordinates, GM, C_R , and C_a

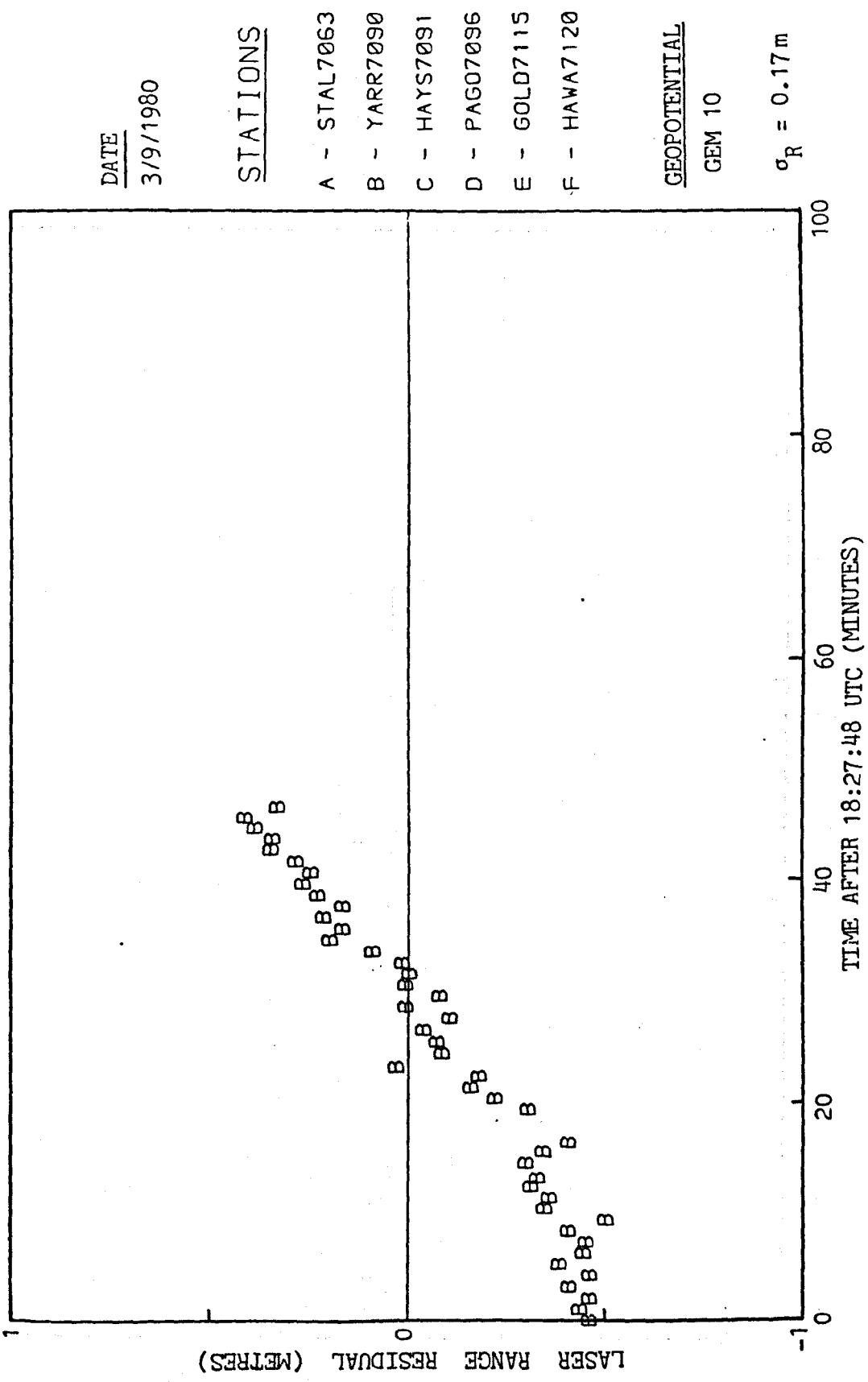


Fig 4.IX Range Residuals From Solution for Satellite State Vector, Station Coordinates, GM, C_R , and C_a

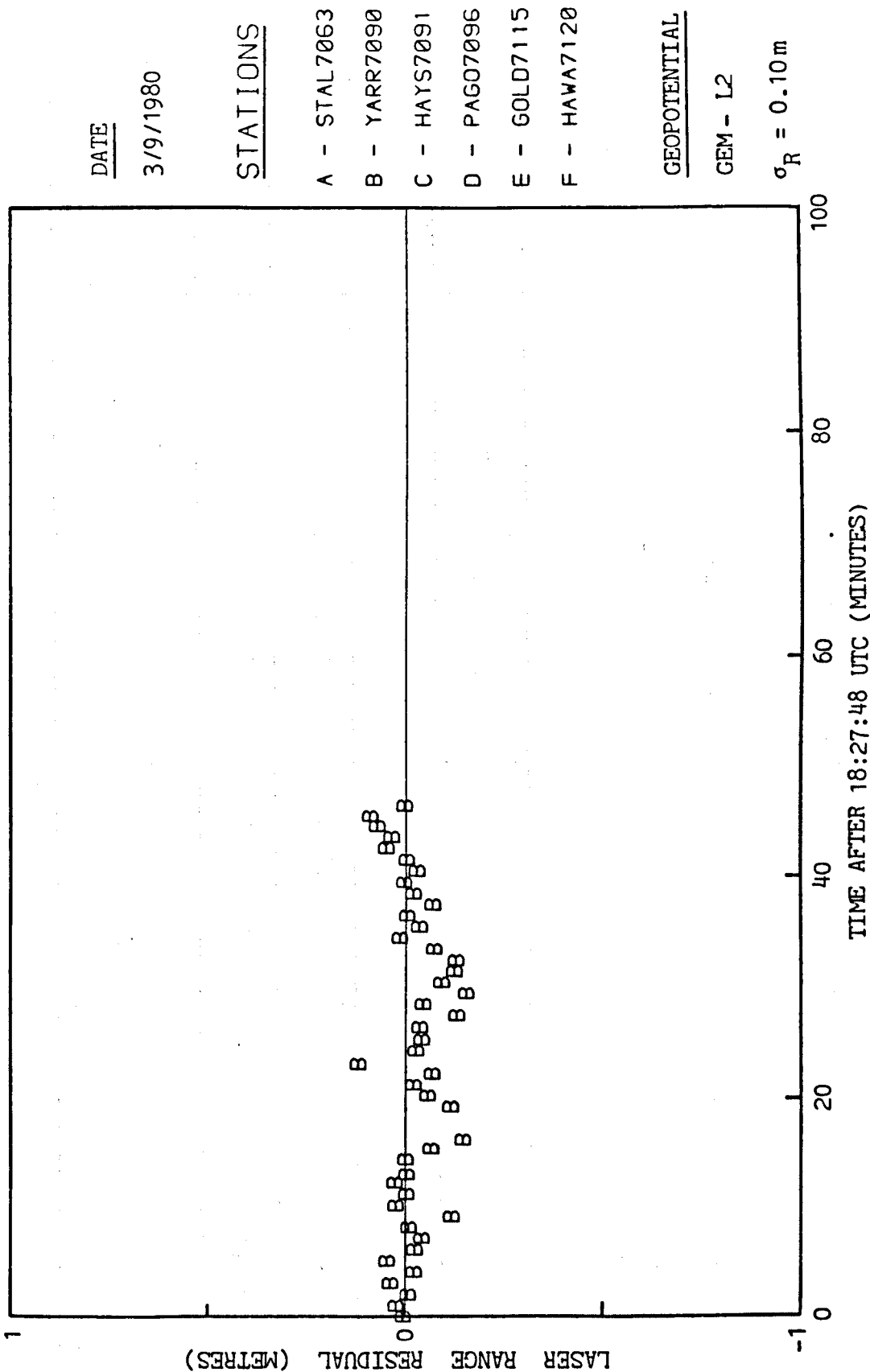


Fig 4.X Range Residuals From Solution for Satellite State Vector, Station Coordinates, C_M , C_R , and C_a

PARAMETER	DS 1		DS 2	
	VALUE	STANDARD ERROR	VALUE	STANDARD ERROR
$C_M^{**} \text{ (Km}^3 \text{ s}^{-2} \text{)}$	398600.605	0.004	398600.605	0.008
$C_M^* \text{ (Km}^3 \text{ s}^{-2} \text{)}$	398600.449	0.003	398600.440	0.004
C_R^*	1.109	0.015	1.110	0.016
$C_a^* \text{ (ms}^{-2} \text{)}$	-8×10^{-12}	2×10^{-12}	5×10^{-12}	3×10^{-12}

NOTES

* : Adjustment used GEM-L2 and c = 299792458 ms⁻¹

** : Adjustment used GEM10 and c = 299792500 ms⁻¹

Fig 4.XI Force Model Parameter Solution

researchers. Lerch et al (1983) obtained $398600.607 \text{ Km}^3 \text{ s}^{-2}$ with a speed of light of $299792500 \text{ ms}^{-1}$ in their derivation of GEM-L2. The recently adopted speed of light of $299792458 \text{ ms}^{-1}$ (Kaplan, 1981) has now been established as a fundamental quantity. This implies a different value of GM (see fig 4.XI), and the MERIT Campaign has adopted $398600.448 \text{ Km}^3 \text{ s}^{-2}$ (Melbourne, 1983). This is very close to the solutions given in fig 4.XI, which are also confirmed by Christodoulidis et al (1982) and Lerch et al (1983) from adjustments with LAGEOS data. With the new speed of light, holding GM fixed at 398600.448 did not seem to affect the solutions or the residuals. In all subsequent adjustments GM was fixed at this value because it was not desirable to solve for it, as this would introduce slight changes in the scale of the network.

The values of C_R and C_a in fig 4.XI are of the same order of magnitude as those obtained by other researchers (Christodoulidis et al, 1983 and Reigber et al, 1982). The solar radiation reflectance coefficient, C_R , is better determined than C_a , because the acceleration arising from solar radiation is three orders of magnitude bigger than the along-track acceleration (see fig 3.III). C_a is thus not an important parameter in 4-day arcs but, since its effect increases as the square of the time (10m after 30 days), it becomes significant with longer arcs.

4.2.3 Solutions for Tracking Station Coordinates and Satellite

Starting Elements

The final solutions for the EF tracking station coordinates from the two data sets are tabulated in fig 4.XII. These values were derived in an adjustment which also solved for satellite starting elements, C_R and C_a . The geopotential model was GEM-L2.

and the value of GM was $398600.448 \text{ km}^3 \text{ s}^{-2}$. The new value of the speed of light was used in order to scale the ranges (see Section 4.2.2). Polar motion and UT1-UTC were linearly interpolated from the values given in the BIH Annual Report (BIH, 1981). The longitude of station STAL 7063 was held fixed at $283^\circ 10' 19''.800$ (see fig 4.I). After adjustment, the rms residual, σ_R , was 10 cm with both data sets.

Referring to fig 4.XII, with the exception of PAGO 7096 all the standard errors from DS 1 were less than 10 cm. The standard errors of PAGO 7096 were of the order of one metre, and this was because only one satellite pass was observed from that station (see fig 4.III). The precisions of the station coordinates from DS2 were generally worse than those from DS 1, because DS2 consisted of considerably fewer normal points and consequently the unknowns were more weakly determined. The standard errors from the latter data set were at the 20 cm level.

A comparison of the geodetic station coordinates between the DS1, DS2 and SL5 (Christodoulidis et al, 1982) solutions is given in fig 4.XIII. The coordinates of the four common stations in the two data sets were repeatable at the metre level. In addition, the DS1 and DS 2 coordinates were in agreement with the SL5 solutions to within just over one metre. In all cases, the station heights were more consistent than the corresponding longitudes and latitudes. All the heights derived from DS1 and DS2, for the four common stations, were within 10 cm of each other. Furthermore, except for PAGO 7096 (large standard errors), all the differences in station heights in fig 4.XIII were under 30 cm. This seemed to suggest that the larger differences in longitudes and latitudes could be due to errors in the BIH values of polar motion. The next section (4.2.4) describes the tests that were performed in order to isolate the

STATION	DS 1 COORDINATES (METRES)			DS 2 COORDINATES (METRES)		
	X σ_x	Y σ_y	Z σ_z	X σ_x	Y σ_y	Z σ_z
STAL 7063	1130711.92 0.01	-4831370.63 0.03	3994089.66 0.06	1130711.96 0.03	-4831370.82 0.12	3994089.45 0.08
HAYS 7091	1492451.37 0.06	-4457281.92 0.03	4296818.40 0.04			
OWEN 7114				-2410427.22 0.11	-4477803.64 0.07	3838688.51 0.08
GOLD 7115	-2350866.20 0.06	-4655547.05 0.05	3661000.64 0.05	-2350866.50 0.11	-4655547.67 0.26	3660999.80 0.08
PASA 7896				-2493215.75 0.10	-4655229.99 0.07	3565577.48 0.09
YARR 7090	-2389003.54 0.06	5043333.50 0.03	-3078527.41 0.05	-2389003.46 0.20	5043333.49 0.06	-3078527.34 0.09
PAGO 7096	-6100049.85 0.10	-996200.16 1.15	-1568977.20 0.64			
HAWA 7120	-5466003.13 0.03	-2404405.75 0.08	2242230.14 0.05	-5466003.15 0.06	-2404406.62 0.13	2242229.06 0.09

Fig 4.XII Solutions for Tracking Station Coordinates

STATION	DS2 - DS1 (METRES)				SL5 - DS1 (METRES)				SL5 - DS2 (METRES)			
	ΔA	$\Delta \phi$	ΔH		ΔA	$\Delta \phi$	ΔH		ΔA	$\Delta \phi$	ΔH	
STAL 7063	0.00	-0.27	0.02		0.00	-1.14	-0.08		0.00	-0.87	-0.10	
HAYS 7091	-	-	-		-0.07	-1.11	-0.03		-	-	-	
OWEN 7114	-	-	-		-	-	-		0.52	0.06	0.03	
GOLD 7115	0.00	-1.08	0.07		0.33	-1.05	-0.21		0.33	0.03	-0.28	
PASA 7896	-	-	-		-	-	-		0.31	-0.25	-0.13	
YARR 7090	-0.08	0.03	-0.07		0.11	0.68	-0.11		0.19	0.65	-0.04	
PAGO 7096	-	-	-		-0.69	-0.09	-0.38		-	-	-	
HAWA 7120	0.79	-1.17	-0.05		0.67	-0.74	-0.11		-0.12	0.43	-0.06	

Fig 4.XIII Comparison Between DS1, DS2, and SL5 Station Coordinates

ELEMENT	DS 1 EPOCH : 2 nd September 1980, 0 ^h 0 UTC		DS 2 EPOCH : 2 nd December 1980, 0 ^h 0 UTC	
	VALUE	STANDARD ERROR	VALUE	STANDARD ERROR
x_o (m)	-6137454.53	0.08	-1128586.88	0.12
y_o (m)	829424.66	0.12	6078033.39	0.10
z_o (m)	1063688.90	0.06	10647597.92	0.08
\dot{x}_o (ms ⁻¹)	4094.78152	0.00004	3134.98865	0.00010
\dot{y}_o (ms ⁻¹)	3318.99755	0.00005	4243.52425	0.00007
\dot{z}_o (ms ⁻¹)	2119.17796	0.00004	-2102.25439	0.00005

Fig 4.XIV Solutions For LAGEOS Starting Elements

effects of any inaccuracies in the BIH system.

The values of the IF LAGEOS starting element from the two adjustments are given in fig 4.IV. Their precisions were of the same order of magnitude as those of the tracking station coordinates. In general, the elements from DS1 had smaller standard errors than those from DS2, because the greater number of observations in the first data set resulted in a better determined satellite orbit.

4.2.4 Polar Motion Solutions

The rather large differences in station coordinates (longitudes and latitudes) between DS1, DS2 and SL5 (see fig 4.XIII) reflect the limit in the ability of the BIH to maintain the EF reference frame. This is because the DS1 and DS2 coordinates were determined using the BIH values of x_p and y_p , and the coordinate differences of the four common stations in fig 4.XIII represent, to a large extent, the inconsistencies in the BIH values. To overcome this problem new polar motion values were derived. The station coordinates were first held fixed at the SL5 values of fig 4.I, and the observations of DS1 were used in order to solve for new polar motion components. The latter were used, in place of BIH, to solve for a new set of DS1 station coordinates. The four common stations between the two data sets were subsequently held fixed at the new DS1 values, and a set of polar motion parameters was derived with the DS2 normal points. The new polar motion values are tabulated in fig 4.XV, along with the corresponding BIH and GEM-L2 (Lerch et al, 1983) solutions. The standard errors of the polar motion components in the Nottingham solutions were all under 0.001 (3 cm).

			DATE		CHANGE Dec - Sept
			3/9/1980	3/12/1980	
POLAR MOTION VALUES	NOTTINGHAM	x_p^*	-0.0025	0.0049	0.0074
		y_p^*	0.00354	0.00396	0.0042
	GEM-L2	x_p^{**}	-0.0026	0.0045	0.0071
		y_p^{**}	0.00343	0.00378	0.0035
	BIH	x_p^{**}	-0.0017	0.0019	0.0036
		y_p^{**}	0.00317	0.00378	0.0061

NOTES :

- * : 4-day average values
- ** : 5-day average values

Fig 4.XV Polar Motion Solutions

In the maintenance of an EF reference frame, what matters is the ability of a certain method to accurately detect changes in polar motion. Absolute values of polar motion can never be obtained, and if two methods detect the same changes then they can be 'calibrated' to give the same values. Between the Nottingham and the GEM-L2 solutions (see fig 4.XV), the difference in the observed change in x_p was 0."003 (0.09m) and in y_p it was 0."007 (0.21m). The corresponding differences between the BIH and the Nottingham values were 0."038 (1.18m) for x_p and 0."019 (0.59m) for y_p .

The new polar motion values of fig 4.XV were used to re-adjust the tracking station coordinates. The new differences between DS1, DS2 and SL5 coordinates are tabulated in fig 4.XVI. There was a very significant improvement over the solutions which used the BIH data, in fig 4.XIII. Coordinate differences between DS1 and DS2, which were previously at the metre level, were now about 40 cm. Agreement between the new DS1 and the SL5 coordinates was of the order of 20 cm. PAGO 7096 was the only exception, and in this case its large standard errors in the adjustment (see fig 4.XII) could explain the rather sizeable deviations from the SL5 coordinates. The DS2 coordinates were slightly less consistent than DS1, being within 45 cm of SL5, because of the smaller number of observations in this data set.

The results in fig 4.XVI reflect the true accuracy of the station coordinates in the orbit determination process, since any errors from polar motion have been removed. Considering the limitations of the two data sets (see Section 4.2.1), significantly better results can be expected with more observations and with data

STATION	DS 2 - DS 1 (METRES)				SL 5 - DS 1 (METRES)				SL 5 - DS 2 (METRES)			
	ΔA	$\Delta \phi$	ΔH		ΔA	$\Delta \phi$	ΔH		ΔA	$\Delta \phi$	ΔH	
STAL 7063	0.00	-0.03	0.01		0.00	-0.12	-0.08		-0.00	-0.09	-0.09	
HAYS 7091	-	-	-		0.05	-0.15	-0.03		-	-	-	
OWEN 7114	-	-	-		-	-	-		0.37	0.19	0.04	
GOLD 7115	-0.35	-0.16	0.06		-0.15	0.06	-0.22		0.20	0.22	-0.28	
PASA 7896	-	-	-		-	-	-		0.21	-0.15	-0.13	
YARR 7090	-0.35	0.07	-0.08		0.08	-0.18	-0.11		0.43	-0.25	-0.03	
PAGO 7096	-	-	-		-0.84	0.37	-0.39		-	-	-	
HAWA 7120	-0.11	0.06	-0.03		-0.02	-0.03	-0.11		0.09	-0.09	-0.08	

Fig 4.XVI Comparison Between DS1, DS2, and SL5 Station Coordinates
After Polar Motion Solution

from the recently introduced 3rd generation laser tracking facilities. It would not be too optimistic to expect accuracies of 2-5 cm in station coordinates in the near future.

When solving for polar motion, a number of quantities have to be held fixed. As mentioned in Section 2.5.4, the values of x_p and y_p define the orientation of the Z-axis of the EF reference frame with respect to the earth's instantaneous spin axis. It is thus impossible to solve for absolute values of x_p and y_p at the same time as solving for all the station coordinates, because this would result in an ill-conditioned normal matrix. The minimum constraint is that two station latitudes have to be held fixed in addition to the fixed longitude. In practice it is preferable to use externally derived values of polar motion (eg by the BIH), with an appropriately large set of observations, in order to determine the best possible station coordinates. These coordinates are then held fixed, with additional sets of observations, in order to solve for new values of x_p and y_p for each data set. The polar motion values then refer to the mid-epoch of the data set from which they were derived.

More investigations are needed in order to determine the optimum observation span for which polar motion values can be assumed to be constant. The current thinking is that 5-day sets of data are most suitable, and this is the interval used by the BIH. In the analysis of DS1 and DS2 there was no improvement in the solutions between using fixed polar motion values, or interpolating the BIH values to the exact observation epoch. So it appears that x_p and y_p can be taken as constant over the 4-day intervals, and this should also be true for 5-day intervals. However, it would be desirable to have the greatest possible resolution in the polar

motion determination in order to be able to observe any short-term changes. The potential of 12-hourly or even 6-hourly determinations should thus be investigated. Such solutions would require significantly larger numbers of observations than those in the data sets described in Section 4.2.1.

When using observations covering a long time span, it is possible to solve for all the station coordinates and for polar motion. This can be done, as explained for UT1 - UTC in Section 2.5.4, by incorporating a number of, say 5-day, sets of x_p and y_p unknowns, and holding the first set fixed. In a single orbit determination all the station coordinates, the satellite starting elements, and all subsequent sets of 5-day polar motion values can be derived. The solutions for successive sets of x_p and y_p represent changes from the initial fixed values. The latter are necessary in order to define the EF reference frame and they can be the BIH values (initially) or the values derived from a previous orbit determination with an appropriate overlap (say 5 days).

Two methods of obtaining polar motion values have been described. The first method is suitable for quick results, and it uses small sets of tracking data, holding all the station coordinates fixed. The second method is suitable for periodic revisions of the tracking station coordinates, and it uses longer (monthly or yearly) data spans, solving for a number of sets of polar motion values. Both methods are necessary in the maintenance of the EF reference frame, with routine determination of x_p and y_p (first method), and occasional revisions of the station coordinates in order to account for real movements such as plate motions (second method).

4.2.5 Effect of Earth Rotation

Unlike polar motion, absolute values of UT1-UTC do not affect the repeatability of station coordinates in separate orbit determinations. The reason is that one can use any values of UT1-UTC and still get identical station coordinates, as long as changes of UT1-UTC are modelled accurately. As already explained (Section 2.5.4), different values of UT1-UTC only affect the solutions for the satellite starting elements, by transforming the orbit to different inertial frames. It is not therefore sensible to solve for both the satellite state vector and for earth rotation in one adjustment. One such solution was attempted with the normal points of DS1 and, while the station coordinates were well determined (with standard errors similar to those given in fig 4.XII), the satellite starting elements were very weak. The standard errors of the components of the initial satellite position vector were of the order of 1 km, while those of the components of the satellite velocity vector were of the order of 1 ms^{-1} .

Orbit determination is very sensitive to changes in UT1-UTC, which are generally termed as changes in the length of day (l.o.d). These can be derived (see Section 2.5.4) either by fixing the satellite orbit and solving for a number of, say daily, UT1-UTC values, or by fixing an initial value of UT1-UTC and solving for subsequent (daily) values, along with the satellite state vector. The former method was used in the derivation of earth rotation with the observations from DS1 and DS2. Two sets of satellite starting elements were initially determined, using the BIH values of UT1-UTC which were linearly interpolated to the observations epochs, and using the full data sets, DS1 and DS2 (see Section 4.2.3 and fig 4.XIV). The normal points were then divided into 8 smaller

sets, each containing the observations for one day. A new adjustment was performed for each day, solving only for UT1 - UTC and holding everything else fixed at the values that were derived in the two adjustments which used the full data sets. The daily solutions for UT1 - UTC and their standard errors are tabulated in fig 4.XVII. The UT1 - UTC values were differenced to give the changes in the l.o.d. , and these changes are also given in fig 4.XVII.

The changes in the l.o.d. were of the order of 0.002 (1 m) per day (see fig 4.XVII). This implies that UT1 - UTC cannot be assumed to be constant for periods longer than a few hours without a significant loss of accuracy. However, solutions for UT1 - UTC with only a few hours' observations are generally weak, necessitating the use of data covering at least 24-hour intervals. Indeed, most researchers solve for 5-day average earth rotation values (Lerch et al, 1983). Such solutions lead to increased least squares residuals, because 5-day intervals are clearly too long to assume a constant value of UT1 - UTC. To demonstrate this point, a constant UT1 - UTC value was used in an adjustment with DS1. The rms residual, σ_R , increased to 28 cm, from the 10 cm (see Section 4.2.3) that were obtained when using the linearly interpolated BIH values, illustrating a significant loss of accuracy.

From the above discussion it is obvious that a better method for determining earth rotation has to be established. With the more accurate 3rd generation laser tracking facilities it may be possible to have resolutions in the UT1 - UTC determinations of as little as 6 hours. Another approach would be to solve for a drift term, in addition to a constant value of UT1 - UTC, for each

	DATE	UT 1 - UTC (msec)	STANDARD ERROR (msec)	CHANGE IN LENGTH OF DAY (msec)
DS 1	2/9/1980	99.95	0.4	-2.28
	3/9/1980	97.67	0.6	-1.77
	4/9/1980	95.90	0.5	-2.57
	5/9/1980	93.33	0.6	
DS 2	2/12/1980	-127.56	0.4	-2.43
	3/12/1980	-129.99	0.8	-2.79
	4/12/1980	-132.78	0.6	-2.04
	5/12/1980	-134.83	0.7	

Fig 4.XVII Daily Earth Rotation Solutions

DATE	AVERAGE CHANGE IN LENGTH OF DAY		
	NOTTINGHAM* (msec/day)	GEM - L2** (msec/day)	BIH** (msec/day)
3/9/1980	-2.2	-1.9	-2.1
3/12/1980	-2.4	-2.5	-2.4

NOTES:

- * : 3-day average
- ** : 5-day average

Fig 4.XVIII Comparison of Changes in l.o.d. Derived
From Different Sources

day. In a single orbit determination with observations spanning a number of days, daily UT1 - UTC values and daily drift terms can be determined. The satellite starting elements can also be included in the list of unknowns, provided the initial (first day's) UT1-UTC value is held fixed (see Section 2.5.4).

The average changes in l.o.d. over the two data sets, derived from the daily solutions in fig 4.XVII, are tabulated in fig 4.XVIII. These are compared with the corresponding BIH (BIH, 1982) and GEM-L2 (Lerch et al, 1983) values. The agreement between the Nottingham and GEM-L2 solutions was better than 0.3msec/day (0.15m/day). However, contrary to the polar motion comparison (see Section 4.2.4 and fig 4.XV), the Nottingham values showed better agreement with the BIH than with GEM-L2. In this case the changes in l.o.d. were within 0.1msec/day (5 cm/day) of the Nottingham solutions.

4.3 Applications of Precise Orbit Determination

Satellite orbit determination has a great number of potential applications, depending on the satellites being tracked and on the accuracy of the observations. The discussion given here will concentrate mostly on geodetic applications, with an emphasis on the most recent and precise techniques available.

Clearly, the most obvious application of OD is the computation of the satellite ephemeris. The accuracy with which this is required depends on the aims of the individual mission. The precision requirements are particularly stringent for altimetry satellites, like the European Space Agency's ERS - 1 (Dow and Klinkrad, 1982) which is due to be launched in 1988.

The ephemerides of navigation satellites, like Transit and GPS, (see Chapter 5), also need to be computed very accurately. Such determinations are made from data collected by networks of dedicated tracking stations. The most up-to-date observations are used in order to derive predicted ephemerides, which are stored on board the satellites and are subsequently transmitted to the users for instantaneous navigation. Post-mission ephemerides are also computed, and made available to bona fide users, for more precise applications like geodetic positioning.

In order to compute a satellite ephemeris a suitable geopotential model must be used (see Section 2.3.2). The derivation of such models is a very important application of orbit determination. The observations from several satellite missions have been adjusted to determine geopotential coefficients, and a number of models are currently in use. Some of these have been derived by combining the satellite tracking data with surface gravimetry (Lerch et al, 1979) in order to obtain higher accuracies in selected regions of the world. In particular, all the even-numbered GEM models utilized surface gravity data.

Satellite observations have been used in the derivation of atmospheric density models (CIRA, 1972) and in studies of solar radiation. Such investigations have necessitated the deployment of specially suited spacecraft. For example, PAGEOS and DASH 2 (Lucas, 1974) were large balloon-type satellites, with high area-to-mass ratios, which were sensitive to drag and solar radiation forces (see Sections 2.3.7 and 2.3.8).

Another important application of precise orbit determination is in geophysics and geodynamics. Satellites have been used in order to test and verify models for solid earth and ocean tides (Cazenave et al, 1977). In addition, the projected accuracy of the tracking station coordinates with SLR observations to LAGEOS of 2-5 cm (see Section 4.2.4), provides a powerful technique to investigate tectonic plate motions and local crustal movements. The US National Astronautics and Space Administration (NASA) is currently running a Geodynamics Program (NASA, 1981), based largely on SLR observations to LAGEOS and on VLBI. The Program's main objectives are to improve our understanding of the solid earth and particularly of the crustal processes associated with natural hazards and resources, to develop geopotential and geomagnetic models, and to facilitate the establishment of new geodynamics measurement services requiring precise positioning. To achieve these objectives a Crustal Dynamics Project (CDP) is responsible for the development of the SLR and VLBI systems, and for the analysis of the data acquired by them. Laser ranging to LAGEOS is also to be used in a European project, Project WEGENER, to investigate the plate-tectonic activity of the Eastern Mediterranean (Haworth, 1984). The analysis of the SLR observations in all the above projects is performed using the techniques of orbit determination that have been described in this thesis.

Orbit determination is an invaluable process in the definition and maintenance of a conventional terrestrial coordinate reference system. COTES (Mueller, 1983) is a joint IAG/IAU working group which has been set up in order to prepare

a proposal for the establishment of such a system and for the service(s) necessary to maintain it (see Section 2.2.3). The definition of an EF reference frame is implicit in the adopted values of the coordinates of a set of globally distributed 'primary' stations which can, for example, be the satellite tracking stations. This system is maintained by routine determinations of polar motion values, holding the station coordinates fixed, as explained in Section 4.2.4. The current thinking is for additional determinations of the station coordinates, as they become necessary, in order to account for real movements between them. Since there are a number of available techniques for the computation of the polar motion values, the MERIT Campaign will intercompare them and try to remove sources of systematic differences between them.

In the maintenance of the Conventional Terrestrial System (CTS), the determination of earth rotation (UT1-UTC) is also of paramount importance. Although, as discussed in Section 4.2.5, absolute values of UT1-UTC are not required for artificial satellite orbit determination they are essential in other techniques such as VLBI or optical astrometry. These techniques are based on observations of the inertially fixed extra-galactic radio sources and stars, resulting in their respective EF (or CTS) reference frames being defined in relation to the inertial frames that are implicit in the adopted coordinates of the various celestial bodies. In orbit determination, on the other hand, the IF is defined relative to the EF reference frame, and different values of UT1 - UTC simply transform the satellite coordinates to different inertial frames, but they do not affect the solutions of any of the EF unknowns (eg station coordinates).

However, as the ERP's are to be determined by a combination of the most accurate techniques available, namely SLR, LLR and VLBI, it is necessary to remove systematic biases between the inertial frames, as well as the EF reference frames that are inherent in each technique. COTES therefore will also formulate proposals for the establishment of a Conventional Inertial System (CIS).

Satellite orbit determination is the only method that provides geocentric station positioning. This, and the low cost and availability of satellite receivers, has led to the use of Transit Doppler observations for a wide range of geodetic applications. In the UK, for example, Transit observations have resulted in the re-adjustment of the primary triangulation and in the Ordnance Survey Scientific Network OSSN80 (Ashkenazi et al, 1981).

Satellite techniques can provide global, as well as local, geodetic control and they are used extensively in worldwide mapping. A future global geodetic network will probably consist of a number of 'primary' stations, whose CTS coordinates will be determined by the most precise techniques available (SLR, LLR, VLBI). Other less precise techniques, like Transit and GPS, will provide densification of the network by establishing the coordinates of a far greater number of 'secondary' stations, which will make the CTS widely accessible to all potential users.

CHAPTER FIVE

THE GLOBAL POSITIONING SYSTEM

5.1 Introduction

The Navigation Satellite Timing and Ranging Global Positioning System (NAVSTAR GPS) is a new satellite navigation system which is currently under development by the US Department of Defence (DOD). GPS will succeed the present satellite navigation system, Transit (Gough, 1978 and Sykes, 1979), and should be fully operational by late 1988. It is designed to provide instantaneous, highly accurate position, velocity, and time information to users anywhere in the world.

The GPS development started after a decision by the US Defence Systems Acquisition Review Council (DSARC) in 1973. Prior to this, two separate projects, the Navy's TIMATION Program and the Air Force Project 621B, had been undertaken. With the decision to deploy GPS these two programs were merged and a Joint Program Office (JPO) was established. The Air Force was designated to lead GPS through its phases, which would result in a single system to serve the vast majority of defence positioning and navigation needs (Parkinson and Gilbert, 1983).

GPS consists of 3 segments (see Section 5.2); the Space Segment, the Control Segment and the User Segment. The Space Segment comprises the satellites, which transmit signals and coded navigation information (see Section 5.3). The original proposal was for the deployment of 24 satellites in 3 orbital planes but this has since been changed to 18 satellites in 6 orbital planes. The User Segment consists of an infinite number of properly equipped users who are able to receive the satellite signals and perform real-time navigation. The Control Segment is responsible for the tracking, orbit determination, upload, and monitoring

functions needed to control the satellites and to provide a continuous and reliable service to the users.

The GPS development program is in three phases, which will lead to an operational system in 1988. Phase I was termed the Concept Validation Program, and took place between 1973 and 1979. Phase II, the Full Scale Engineering Development and System Test, requires the maintenance of a 5-satellite constellation in order to allow maximum coverage at the Army Proving Grounds in Yuma, Arizona. This phase includes the development of a prototype operational satellite, an operational Control Segment, and prototype user equipment. The satellite build-up in Phase III, the Full Operational Capability, will start in 1986, when the 5-satellite constellation will very quickly be extended to the operational 18-satellite configuration with an additional 3 on-orbit spares. As GPS becomes operational, the DoD will cease supporting other navigation systems. Transit, in particular, will be phased out by 1992.

GPS will be available to civil, as well as military users. Currently there are provisions for two navigation services. The Precise Positioning Service (PPS) will utilise the P code (see Section 5.3.3) in order to provide the highest attainable accuracies, but will probably be restricted to military users. The Standard Positioning Service (SPS) will provide degraded accuracies (of about 100m) using the C/A code (see Section 5.3.2) and will be made available worldwide to civilian users (Porter et al, 1984). It is envisaged that an annual user charge may have to be paid in order to use either of the above services.

This chapter describes the GPS segments (Section 5.2). It also explains in detail the GPS signals (Section 5.3) and outlines the theoretical background behind the use of GPS for both instantaneous navigation and for high precision geodetic applications (Section 5.4). Section 5.5 gives details of some receiver systems that are currently under development.

5.2 GENERAL SYSTEM DESCRIPTION

5.2.1 Space Segment

The operational GPS Space Segment (after 1988) will consist of 18 satellites in 6 orbital planes inclined at 55° to the equator. The satellites will be placed at altitudes of approximately 20180 Km and will travel in circular orbits with a period of half a sidereal day ($11^{\text{h}}58^{\text{m}}$). This implies that each satellite will repeat exactly the same ground track every two revolutions, thus ensuring that it is visible from a single control station (see Section 5.2.2) at least once a day.

The 6 orbital planes will be equispaced around the equator and the satellites will be placed at 120° intervals in each plane. In addition to the 18 operational satellites there will be 3 active on-orbit spares which will be placed in every other orbital plane and will be manoeuvred to replace any satellites that fail. The phasing of the satellites from plane to plane will be such that each satellite will be 40° ahead of one in the plane directly to the west and 40° behind one directly to the east (see fig 5.1).

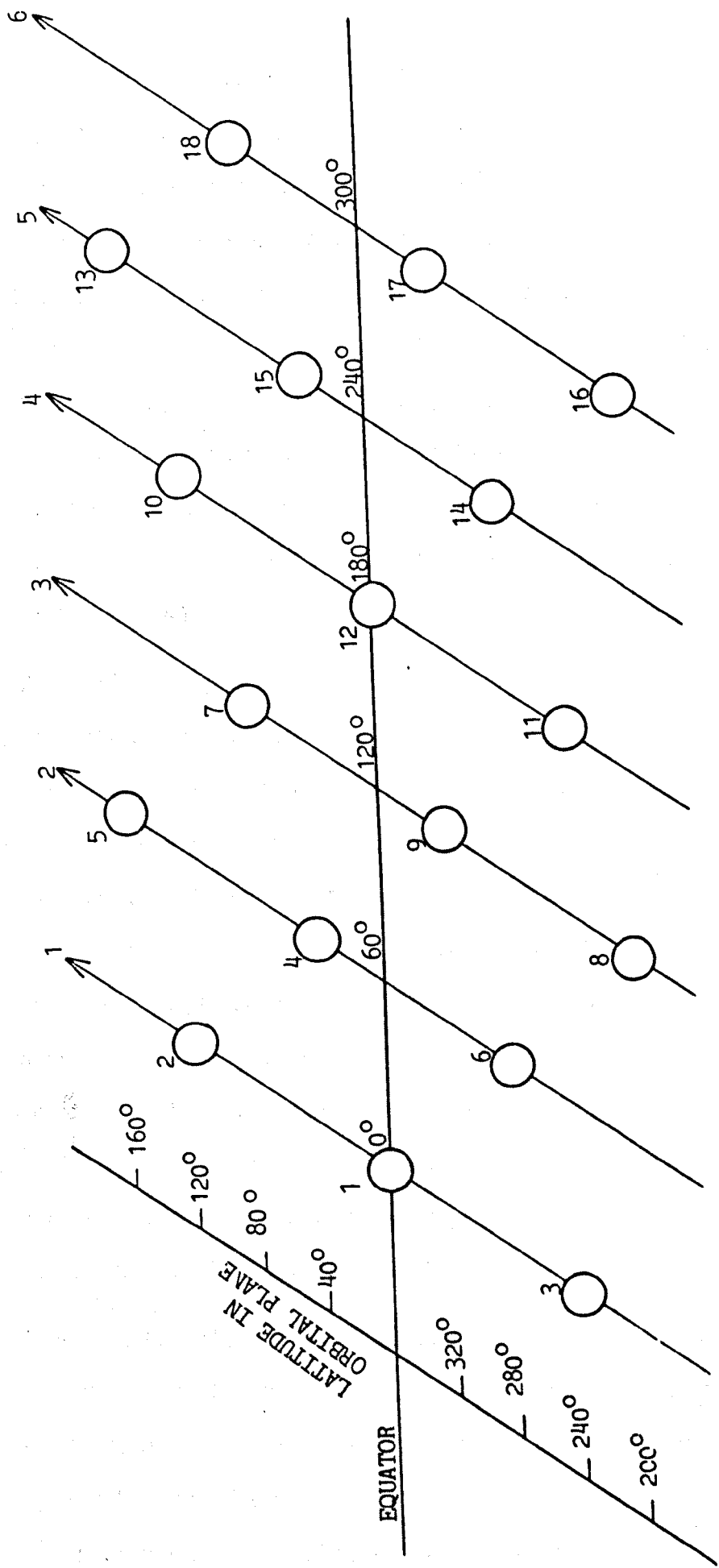


Fig 5.I Operational GPS Satellite Configuration

The satellites transmit signals on two L-band radio frequencies (RF): L1 at 1575.42 MHz and L2 at 1227.60 MHz. These signals carry modulations from two pseudo-random noise (PRN) codes, the P (Precision) code at 10.23 MHz and the C/A (Coarse/Acquisition) code at 1.023 MHz. The PRN ranging codes (see Sections 5.3.2 and 5.3.3) provide specific timing marks on the carrier frequencies and enable the GPS user equipment to carry out the pseudo-range measurements by recording the time of arrival of identifiable satellite signals. The GPS codes are modulated with a data message at a rate of 50 Hz, which supplies the users with essential ephemeris, timing, and almanac information (see Section 5.3.4).

The GPS navigation technique (see Section 5.4.1) consists of making simultaneous pseudo-range measurements to 4 satellites and formulating observation equations for the 4 unknowns of the receiver antenna geocentric coordinates and the clock bias from GPS time. For this technique to work, at least 4 satellites must always be in view anywhere in the world. The configuration of the operational satellites (fig 5.1) provides for global 4-satellite coverage, but in certain regions there will be periodic (once or twice daily) degradations of the 3-dimensional accuracy arising from poor geometry (Porter et al, 1984). Navigation solutions will still be possible but in some cases it may be preferable to perform 2-dimensional solutions by providing a height input to the GPS receivers.

The operational satellite payload will consist mainly of atomic frequency standards for accurate timing (stability of 2-5 parts in 10^{13} per day), a processor to store navigation data, and a PRN signal assembly and L-band antenna for generating and broadcasting the GPS signals. Electrical power will be provided by two solar energy

converting panels that will continually track the sun and charge three batteries for use when the satellites cross the earth's shadow. The Space Vehicles (SV's) will have a mission duration of 6 years and a design life of 7.5 years.

The present (Phase II) satellite constellation of 5 SV's has been designed for maximum 4-satellite coverage at the Yuma Proving Grounds. These satellites are at an inclination of 63° instead of the 55° specified for the operational Space Segment, due to constraints imposed by the Atlas launch vehicle. Future launches on board the Space Shuttle will place satellites at 55° inclinations. The initial Phase III Space Segment will thus be made up of satellites at both inclinations. The build-up to the operational GPS constellation of 18 satellites and 3 spares will start in late 1986 and is expected to finish by the latter part of 1988.

5.2.2 Control Segment

The GPS Control Segment consists of : (a) a number of widely distributed Monitor Stations (MS's) which passively track the satellites, (b) a Master Control Station (MCS) which uses the MS data to compute satellite ephemeris and clock predictions, and (c) a number of Ground Antennas (GA's) which upload these predictions as well as any other Control Segment commands to the satellites and receive telemetry data from them.

The Operational Control System (OCS) will be responsible for initiating sufficient satellite uploads, in order to ensure that the satellite timing and ephemeris prediction user range error (URE) does not exceed 6m (Payne, 1982). The MCS will be located at the Consolidated Space Operations Center (CSOC) at Colorado Springs,

Colorado. There are plans for five MS's, at Ascension Island, Diego Garcia, Kwajalein, Colorado Springs and Hawaii (see fig 5.II). The GA's will be located at 3 of the MS sites. Satellite uploads will be performed at least three times a day per satellite, in order to update the satellite ephemerides. The present (Phase II) Initial Control System (ICS) has 4 MS's (see fig 5.II), at Guam, Hawaii, Alaska and California. The MCS and single GA are both located at Vandenberg Air Force Base, California.

The computation of each satellite's ephemeris is currently performed in two stages (Varnum and Chaffee, 1982) : in the first stage the Naval Surface Weapons Center (NSWC) derives a reference trajectory spanning 40 days, based on one week's tracking data, using the CELEST computer program (O'Toole, 1976). The reference trajectories (satellite ephemerides and associated partial derivatives) are posted to the MCS every 2 weeks. In the second stage the MCS uses the reference orbits, along with the most up-to-date tracking data, in a sequential least squares algorithm (Kalman filter) in order to derive improved satellite ephemerides and clock correction polynomials. These are immediately uploaded to the satellites via the Ground Antenna. The above process follows the same principles as the Nottingham orbit determination software described in Chapter 3. Here too, the orbit determination is performed in two stages, through an orbit integration program (ORBIT), followed by a least squares adjustment procedure (SOAP). The GPS ephemeris is uploaded to each satellite as soon as possible after it becomes visible at Vandenberg, in order to allow the maximum time for testing at the Yuma Proving Grounds.

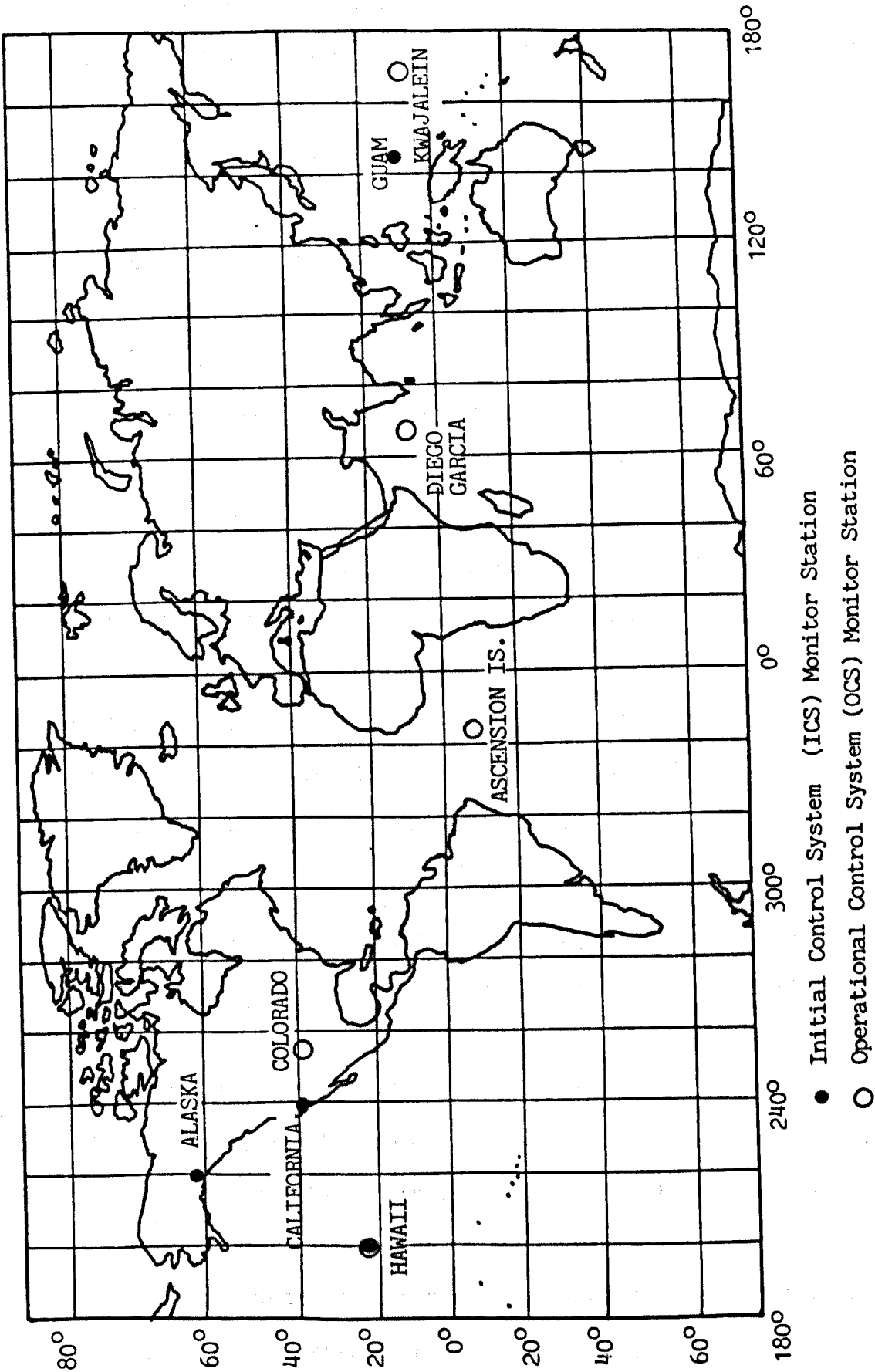


Fig 5.II ICS and OCS Monitor Stations

The MCS orbit determination solves for the six satellite starting elements, two polar motion parameters, two solar radiation pressure unknowns per satellite (see Section 2.3.8), and three satellite clock states. The latter represent a clock drift, bias, and aging term of the satellite with respect to the MCS clock. Two clock unknowns for each of the MS's are also determined, but the MCS clock is held fixed as it is the time base for the GPS system.

5.2.3 User Segment

GPS can support an infinite number of properly equipped users. The task of the satellite receivers is to process the GPS signals according to the users' requirements. Several types of equipment are being developed for a variety of applications and receivers have been built for High Dynamic (HD), Medium Dynamic (MD) and Low Dynamic (LD) environments. These are capable of using the GPS signals in order to perform instantaneous navigation. High precision geodetic receiver systems have also been developed, with the facility to record and post-process the GPS data (see Section 5.5).

The receivers are generally composed of an antenna, a pre-amplifier, a receiver-processor unit (RPU), and a control-display unit (CDU). They vary in complexity and cost according to the specified accuracy and the environment for which they are designed to operate. Multi-channel sets, suitable for HD environments, are the most expensive and complex. The receivers become simpler with fewer channels, and the most basic are the single-channel single-frequency sets. A recent design innovation is for one receiver channel to quickly multiplex up to four satellites in both frequencies (Ward, 1982), thus emulating a multi-channel set.

For a 3-dimensional instantaneous navigation solution, four simultaneous pseudo-range observations are needed. These are used to form observation equations to enable the receiver processor to solve for the antenna WGS-72 coordinates and for the local clock bias from GPS time. To form the observation equations the receiver needs to compute the satellite coordinates from the Keplerian ephemeris representation contained in the navigation message (see Section 5.3.4). Thus a receiver has to acquire and decode the satellite navigation message, as well as to perform the navigation solution.

The GPS WGS-72 system is an earth-centred earth-fixed coordinate system which is implicitly defined by the adopted MS coordinates, gravity field, and earth rotation parameters (see Section 2.5.4). It has the unique advantage of being a global and universally available (through GPS) reference system which should revolutionize navigation, mapping, and geodesy in the not too distant future, provided that the full GPS accuracy is released to all users.

5.3 GPS SATELLITE SIGNALS

5.3.1 Signal Structure and Receiver Measurement Sequence

The satellites transmit signals centred on two L-band carrier frequencies, L1 at 1575.42 MHz and L2 at 1227.60 MHz. Both frequencies are modulated by a PRN code at a chipping rate of $10.23 \text{ Mbits s}^{-1} (\text{MHz})$ called the P (Precision) code. This is a one week-long binary sequence which is unique for each satellite. The L1 carrier is also modulated by a second PRN code, the C/A (Coarse/Acquisition) code, at a rate of 1.023 MHz. This is also a binary sequence, peculiar to each satellite, which is 1023 bits long and repeats itself every millisecond. Both frequencies and codes are derived from a single

on-board 10.23 MHz oscillator. The L1 and L2 carriers are integer multiples of this frequency, giving :

$$f_{L1} = 1575.42 \text{ MHz} = 154 \times 10.23 \text{ MHz} \quad (5.1)$$

$$f_{L2} = 1227.60 \text{ MHz} = 120 \times 10.23 \text{ MHz} \quad (5.2)$$

The reason for the two-frequency transmissions is to enable the users to correct for first order ionospheric effects (see Section 5.4.5.1). Before modulating the carrier frequencies, both PRN codes are modulo - 2 added to a 50 bps (bits per second) data message (see Section 5.3.4). The latter carries satellite ephemeris, almanac, and timing information which the users require in order to perform a navigation solution.

The L1 signal transmitted by the j^{th} satellite at time, t , is

$$\begin{aligned} S_{L1j}(t) = & A_{L1j} P_j(t) D_j(t) \cos 2\pi f_{L1} t \\ & + B_{L1j} C_j(t) D_j(t) \sin 2\pi f_{L1} t \end{aligned} \quad (5.3)$$

In the above equation, A_{L1j} and B_{L1j} represent the amplitudes of the two signal components, $P_j(t)$ and $C_j(t)$ represent the ± 1 P and C/A code sequences respectively, $D_j(t)$ represents the ± 1 data modulations, and f_{L1} the frequency of the L1 carrier. From (5.3) it is seen that the phase of the P signal component leads that of the C/A signal by 90° . The L2 carrier at time, t , is given by

$$S_{L2j} = A_{L2j} P_j(t) D_j(t) \cos 2\pi f_{L2} t \quad (5.4)$$

In this case there are no C/A code modulations, although the capability exists to transmit them, instead of the P code, if required.

The PRN ± 1 modulations take the form of 180° phase reversals of the carrier frequencies. These are illustrated in fig 5.III. Diagram (a) shows a PRN code sequence which has to modulate the 'clean' carrier of diagram (b). The resulting signal is shown in fig 5.III(c).

The power of the transmitted satellite signals is spread over a frequency band of 20.46 MHz ($=2 \times 10.23$ MHz) for the P code and of 2.046 MHz for the C/A code. A typical signal power spectral density plot for L1 (Spilker, 1980) is shown in fig 5.IV. This 'spread-spectrum' technique results in weak signals, with power levels below that of the thermal noise of the receiver. To perform pseudo-range measurements a GPS receiver performs a cross-correlation operation between the received PRN code and a locally generated replica code. This operation concentrates most of the power of the spread-spectrum signals from the specified satellite into the 100 Hz frequency band of the data message, thus vastly improving the signal-to-noise ratio. The signals from other satellites are further suppressed and interference effects are therefore minimized.

The receiver first acquires a signal from a given satellite by cross-correlating the incoming C/A code with a replica code generated by the local (receiver) oscillator. The replica C/A code is shifted in time until it exactly matches the satellite code, and the amount of shift is an ambiguous (by $1\text{ ms} = 300\text{ km}$) measure of the pseudo-range in units of time. The effect of this correlation is to demodulate the satellite carrier from the C/A code. Equation (5.3) then becomes

$$S_{L1j} = C'_j(t) A_{L1j} P_j(t) D_j(t) \cos 2\pi f_{L1} t + B_{L1j} D_j(t) \sin 2\pi f_{L1} t$$

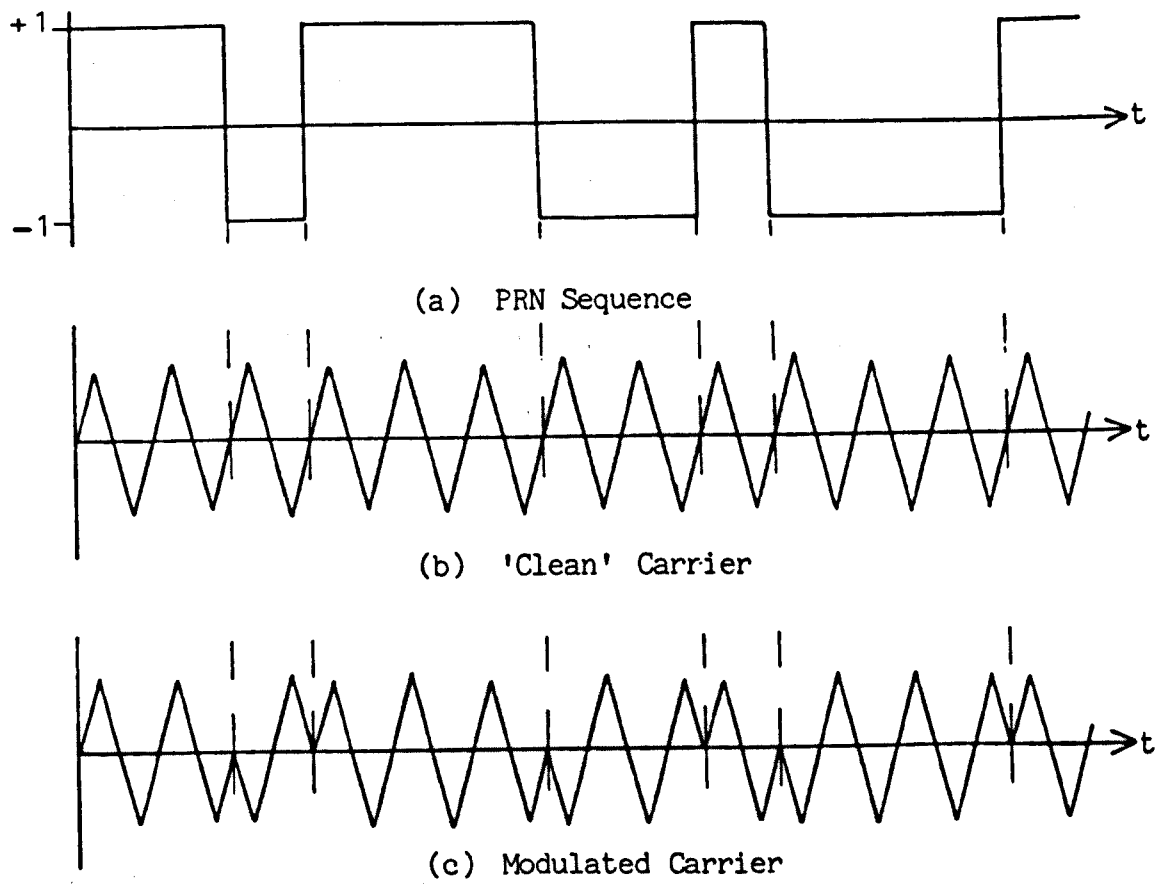


Fig 5.III PRN Modulations

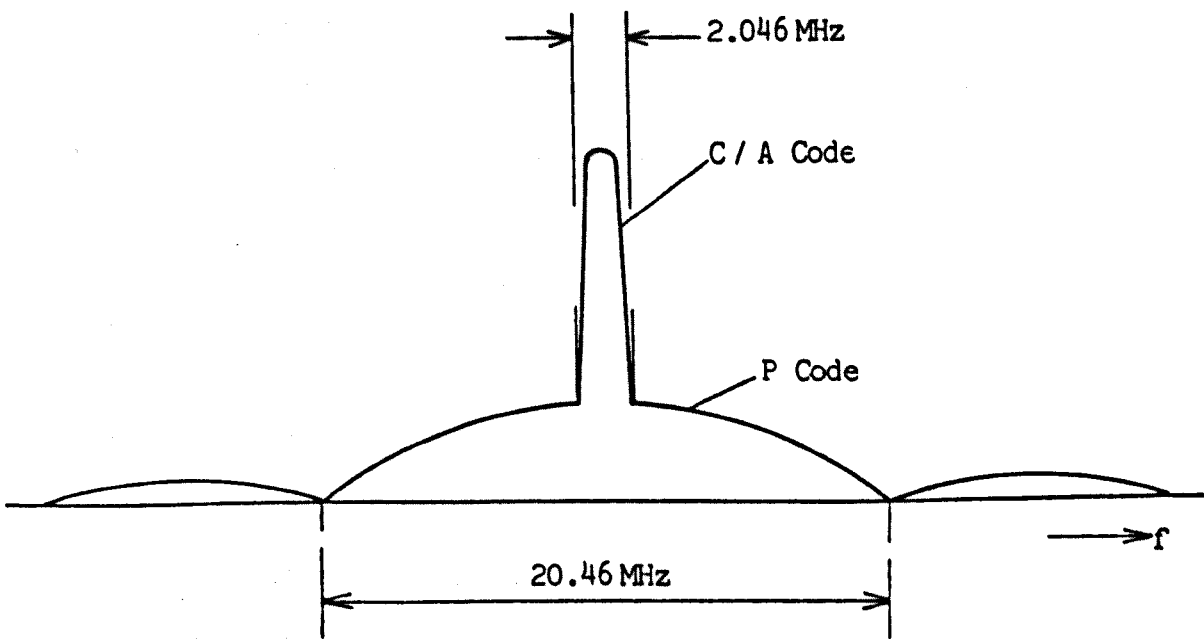


Fig 5.IV L1 Signal Power Spectral Density

where $C_j'(t)$ is the locally generated C/A code. The original ± 1 C/A modulations become $+1$ as a result of the multiplication of the two identical codes. The P code component and signals from other satellites are further suppressed, while the receiver gains access to the satellite data message, $D_j(t)$, on the C/A carrier. This provides it with enough information (see Section 5.3.4) in order to generate the relevant portion of the replica P code and repeat exactly the same procedure as for the C/A code in order to make precise and unambiguous pseudo-range measurements.

The above procedure is necessary for quick initial acquisition of the P code. A P code search without the relevant data message information would be a long process, since this code is one week long (6.37×10^{12} bits). In contrast, the length of the C/A code is only 1023 bits. Any available satellite almanac information will facilitate the starting operation by enabling the receiver to determine which satellites are visible and by making it possible to predict the doppler shift of the satellite signals.

After initial acquisition the receiver can continue tracking the P code, making pseudo-range measurements at specified (local or satellite) time intervals. The data message is available on both L1 carrier components, enabling the receiver processor to compute the satellite coordinates and clock correction polynomials and to perform the navigation solution.

The cross-correlation process demodulates the satellite carrier frequency and results in a 'clean' (reconstructed) sinusoid. The cycles of the beat frequency between this reconstructed carrier and a locally generated signal can be counted, giving a very precise Doppler measurement (see Section 5.4.3) with a resolution of a

fraction of the carrier wavelength ($\lambda_{L1} = 19\text{ cm}$). A Doppler (phase) measurement can alternatively be obtained by measuring the shift of the receiver replica code necessary to enable it to keep track of the satellite code, after initial lock-on.

5.3.2 The C/A Code

The PRN codes of the Global Positioning System are binary bit trains which modulate the carrier frequencies. A binary sequence is normally regarded as a series of 1's and 0's. The basic operation in this discussion is modulo-2 addition of binary numbers, defined by:

$$\begin{aligned} 1 \oplus 1 &= 0 \\ 1 \oplus 1 \oplus 1 &= 1 \end{aligned} \tag{5.5}$$

As mentioned in the previous section, the PRN modulations of the GPS carriers take values of ± 1 . In this case the $+1$ is equivalent to binary 0 and the -1 is equivalent to 1. The operation that corresponds to modulo-2 addition is multiplication, and equations (5.5) become

$$\begin{aligned} -1 \times -1 &= 1 \\ -1 \times -1 \times -1 &= -1 \end{aligned} \tag{5.6}$$

The P and C/A codes are produced by the modulo-2 addition of the outputs of a number of linear feedback shift registers (LFSR's). Before going on to explain the formation of the C/A code it is appropriate to outline the main characteristics of LFSR's.

An n -stage LFSR can be represented by an array of binary numbers with n locations. Fig 5.V illustrates the operation of a 4-stage LFSR. At each clock pulse the state of the array changes, so that all the numbers shift forward by one location. The output of the LFSR is the number occupying the last location of the previous state. As all the numbers shift forward, the first location has to be filled by a new number. This number is derived by the modulo-2 sum of the numbers occupying specified positions in the previous state. The 4-stage LFSR in fig 5.V has tapping points at locations 3 and 4 and its initial (0) state is 0 1 0 0. In the next (1) state the output of the register is the last 0, while the first location is filled by the modulo-2 sum of locations 3 and 4 of the initial (0) state. This process is repeated at every clock pulse, resulting in the PRN output shown in fig 5.V.

When the taps are properly assigned a LFSR cycles through $2^n - 1$ states. The PRN sequence at the output of the 4-state register in fig 5.V thus repeats itself every 15 bits. It can be seen that the 15th state is identical to the initial one and that the LFSR cycles through all possible state vectors except the all-zero state.

The properties of LFSR sequences are discussed in detail in Spilker (1978). Briefly, if a sequence is modulo-2 added to itself (autocorrelation) the result is all zeros (all ones in the ± 1 representation). If a sequence is added to a delayed version of an identical sequence the result is another shifted version of the same sequence.

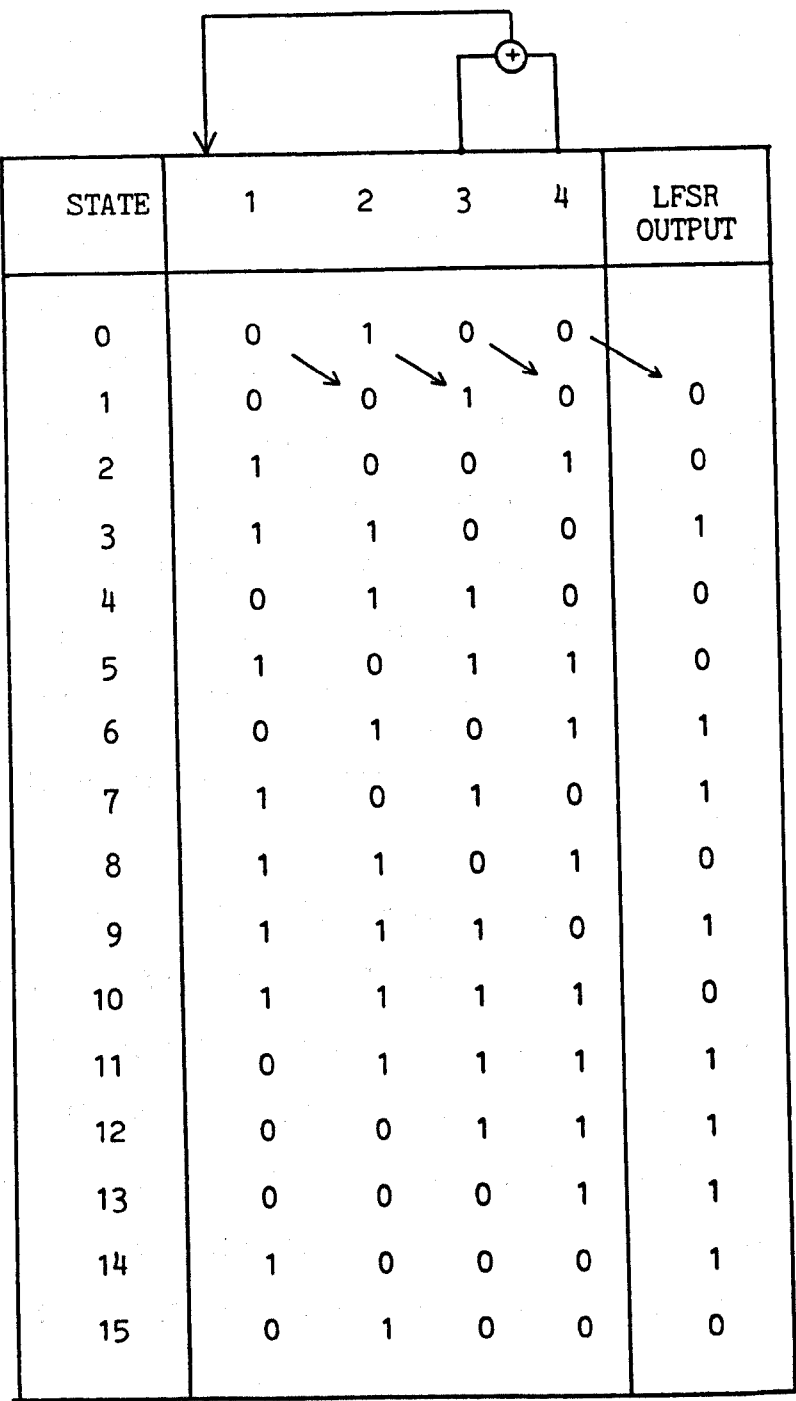


Fig 5.V 4-Stage Linear Feedback Shift Register

The C/A codes ($C_j(t)$ in equation (5.3)) are a family of Gold codes (Gold, 1967) formed as the modulo-2 sum of two binary sequences, G1 and G2. The latter represent the output of two 10-stage LFSR's. The length of the C/A code is thus 1023 ($=2^{10}-1$) bits. There are 1025 possible C/A codes. Of these, 1023 are formed by delaying the G2 sequence, relative to G1, by one bit at a time. The remaining 2 are the G1 and G2 sequences themselves. Currently, 36 different C/A codes have been assigned in the GPS documentation (ICD-GPS-200, 1981). They are formed by having different initial conditions for the G1 and G2 registers. In practice, the G1 state vector consists of all ones, while the state vector of the G2 register varies according to the Space Vehicle identity (SV ID) so that each satellite transmits its own unique code.

A GPS receiver has to have prior knowledge of which satellite transmits a particular code. In other words it must be able to generate all available codes in the GPS system. This enables the receiver to identify the satellites whose signals it is processing. The initial acquisition of the satellite signals is facilitated by the use of the almanac information which is transmitted in the satellite data message (see Section 5.3.4). By knowing which satellites are visible, the receiver can limit the C/A code search (see Section 5.3.1) to the ones which provide the best geometry.

The C/A code is transmitted at a chipping rate (frequency) of $1.023 \text{ Mbits s}^{-1}$ (MHz). Since the complete sequence is 1023 bits long, each code repeats itself every millisecond. This implies that the C/A pseudo-range measurements have a 300 Km ambiguity which can easily be resolved by an initial estimate of the receiver's location (e.g. from a map).

5.3.3 The P Code

The P Code ($P_j(t)$ in equations(5.3) and (5.4)) is a PRN ranging code which represents the modulo-2 sum of two binary sequences, X1 and X2, and is transmitted at a chipping rate of $10.23 \text{ Mbits s}^{-1}$. X1 is 15345000 bits long, while X2 is 37 bits longer (15345037). The period of the combined sequence (P code) is

$$T_P = \frac{15345000}{37} \times \frac{15345000}{10230000} = 622094.6 \text{ s} \quad (5.7)$$

The P code is thus slightly over one week long. In practice it is reset at the beginning of each week for all satellites, so that GPS time starts at midnight (UTC) every Saturday.

The X1 sequence is the modulo-2 sum of the output of two 12-stage LFSR's, X1A and X1B. These are short-cycled to 4092 and 4093 bits respectively. The X1 sequence is reset when X1A completes 3750 cycles, that is to say after 15345000 ($= 3750 \times 4092$) bits. At the 10.23 MHz chipping rate the X1 sequence repeats itself every 1.5 seconds. The completion of a 1.5 second X1 cycle is termed the X1 epoch or Z-count, and there are 403200 such epochs in one week. The Z-count is transmitted in the satellite data message (see next section) and assists the receiver in the initial acquisition of the P code.

The X2 sequence is the modulo-2 sum of a further two 12-state LFSR's, X2A and X2B, short-cycled to 4092 and 4093 bits respectively. The sequence is reset after 3750 cycles of X2A, but at the end of each 1.5-second cycle the two LFSR's are forced to remain at their final states for 37 additional pulses. This results in an X2 sequence that is 37 bits longer than X1.

At the start of the week, X1 and X2 begin their cycles together. On completion of the first X2 epoch, X1 has already generated an additional 37 bits. On completion of the second X2 epoch, X1 has generated 2×37 bits, and so on until just over one week (equation (5.7)) when the two sequences start their cycles together again. However, the P code is not allowed to run to its full length and it is reset at midnight (UTC) every Saturday.

There are 37 mutually exclusive P codes, formed by delaying the X2 sequence by from 1 to 37 bits respectively, before the modulo-2 addition with X1. They are assigned to each satellite according to the SV ID, so that each one transmits its own unique P code. Details of the present code generation and assignment to individual SV's are given in ICD-GPS-200 (1981).

5.3.4 Satellite Data Message

The data message, $D_j(t)$, is a binary bit stream at a rate of 50 bits s^{-1} which is modulo-2 added to the P and C/A codes before they are used to modulate the GPS carriers. A complete data frame (1500 bits) consists of 5 subframes of 300 bits and is transmitted in 30 seconds. Subframes 4 and 5 are each divided into 25 pages, so the complete navigation message is transmitted in 12.5 minutes. The message includes satellite ephemeris and clock correction information, as well as almanac and health data for all the GPS satellites. The almanac enables the receiver to establish the identities of all visible SV's and to select those which provide the best geometry for a navigation solution.

Each data subframe consists of ten words of 30 bits. The first two words are always a telemetry (TLM) and a handover (HOW) word. The TLM word is transmitted first and contains information that is used by the Control Segment. The HOW word is necessary to facilitate the initial acquisition of the P code by the GPS receiver. It enables the receiver to switch from the C/A to the P code after receiving only one subframe (6 seconds) of data (see sections 5.3.1 and 5.3.3). The HOW word contains a truncated version of the Z-count (see Section 5.3.3) called the time-of-week (TOW) count. The TOW count, if multiplied by 4, gives the Z-count which occurs at the start of the next data subframe. It thus enables the GPS receiver to generate the appropriate portion of the P code and thus to quickly achieve the initial cross-correlation with the incoming satellite P code. The relationship between satellite time, t_s , in seconds, TOW, and Z-count is given as:

$$t_s = 6 \times \text{TOW} = 1.5 \times \text{Z-count} \quad (5.8)$$

Since GPS time is reset every week, the TOW is a number ranging from 0 to 100799, while the Z-count ranges from 0 to 403199.

All the satellite ephemeris and timing information is transmitted in data subframes 1, 2 and 3. This means that after initial acquisition of the satellite signals, the GPS receiver need only spend 30 seconds per satellite (equivalent to one full frame) before it receives enough data for a position determination.

The satellite clock correction parameters are contained in data subframe 1. These are computed by the Control Segment as part of the MCS orbit determination. The algorithms for correcting the satellite time of transmission to GPS time are outlined in Appendix G.1.

Data subframes 2 and 3 contain the broadcast satellite ephemeris. This is transmitted in the form of Keplerian - type orbital elements. Each set of elements is derived by the MCS, using a least squares fit to 4 hours of satellite cartesian coordinates from the output of the orbit determination. The elements are therefore valid for 4 hours, but they are updated every hour in order to ensure that the users receive the most recent values. Each set of elements is thus valid for the period of its transmission and for 3 hours thereafter. The fit used in the generation of the orbital elements ensures a user range error (URE) of less than 0.4 metres (one sigma). This means that the range error involved in using the Keplerian representation rather than the actual satellite cartesian coordinates (from the orbit determination) will not exceed 0.4 metres. The procedure for computing the WGS-72 cartesian coordinates from the orbital elements is outlined in Appendix G.2.

Data subframes 4 and 5, which are transmitted in 25 pages, mainly consist of truncated ephemeris parameters (almanac) for all the available GPS satellites. These enable the GPS receiver to decide on which satellites are visible at the time of observation and which of these provide the best geometry for a navigation solution.

Other parameters which are transmitted in the navigation message include flags which give details on the health and accuracy of the various satellites and which enable the receiver to calculate parity for the data. In addition, there are parameters for correcting ionospheric delays (see Section 5.4.5.1), which can be useful for the single-frequency users. The relation-

ship between GPS time as kept by the MCS and UTC is also included in the navigation message, and this is very important in time-transfer applications of GPS.

5.4 MODES OF OBSERVATION AND ADJUSTMENT FOR PRECISE POSITIONING

5.4.1 Instantaneous Navigation Principles

This research is aimed primarily at the most precise applications of GPS. However, before going on to describe the models used for precise positioning, this section will give a brief account of the basic principles behind real-time navigation.

A GPS navigation receiver has to make at least four quasi-simultaneous pseudo-range measurements to different satellites. This can be achieved either by having 4 or more receiver channels, or by using a receiver which can quickly sequence (multiplex) a minimum of 4 satellites. The pseudo-range measurement to the j^{th} satellite is defined as

$$\rho_j^{o'} = c (\hat{t}_{rj} - t_{sj}) \quad (5.9)$$

where $\rho_j^{o'}$: observed pseudo-range

\hat{t}_{rj} : local clock time of reception of a GPS signal from the j^{th} satellite

t_{sj} : satellite time of transmission of received signal

c : speed of propagation of electromagnetic waves in vacuo (2.99792458×10^8 m/s).

In general, the satellite clock is not in phase with GPS time. Therefore, the satellite time of transmission has to be corrected to the corresponding GPS system time using

$$t_{tj} = t_{sj} - \Delta t_{sj} \quad (5.10)$$

where t_{tj} : GPS time of signal transmission

Δt_{sj} : correction computed from parameters in the
satellite navigation message (see Appendix G.1)

The corrected pseudo-range observation becomes

$$\ell_j^o = c(t_{rj} - t_{tj}) \quad (5.11)$$

giving

$$\ell_j^o = c(t_{rj} - t_{sj} + \Delta t_{sj}) \quad (5.12)$$

The mathematical model for (5.12) is

$$\ell_j = R_j + c\alpha_o + \cancel{E_{ion}} + \cancel{E_{trop}} \quad (5.13)$$

where α_o : clock bias between GPS and local clocks

E_{ion} : (positive) error arising from ionospheric delays
(see Section 5.4.5.1)

E_{trop} : (positive) error arising from tropospheric delays
(see Section 5.4.5.2)

The true range between receiver and satellite is

$$R_j = [(X_j - X)^2 + (Y_j - Y)^2 + (Z_j - Z)^2]^{\frac{1}{2}} \quad (5.14)$$

where X_j, Y_j, Z_j : WGS-72 cartesian coordinates of j^{th} satellite derived from the elements in the navigation message (see Appendix G)

X, Y, Z : WGS-72 (unknown) cartesian coordinates of the phase centre of the receiver antenna

At least four equations similar to (5.13) must be formulated in order to solve for the receiver's WGS-72 coordinates and for the clock bias, α_0 . In practice, equation (5.13) can be linearized about some approximate values of the receiver coordinates and clock bias, resulting in the least squares observation equation (see Section 2.5.2):

$$\frac{\partial \ell_j}{\partial X} \Delta X + \frac{\partial \ell_j}{\partial Y} \Delta Y + \frac{\partial \ell_j}{\partial Z} \Delta Z + \frac{\partial \ell_j}{\partial \alpha_0} \Delta \alpha_0 = (\ell_j^o - \ell_j^c) + v \dots (5.15)$$

where $\Delta X, \Delta Y, \Delta Z$: least squares corrections to the provisional receiver coordinates

$\Delta \alpha_0$: least squares correction to provisional value of the clock bias

v : least squares residual.

The observed pseudo-range, ℓ_j^o , is given in equation (5.12), while the computed pseudo-range is

$$\ell_j^c = R_j^c + c \alpha_0^c + E_{\text{ion}} + E_{\text{trop}} \dots (5.16)$$

where α_0^c : computed (provisional) value of the clock bias

The computed true range, R_j^c , is

$$R_j^c = [(X_j - X^c)^2 + (Y_j - Y^c)^2 + (Z_j - Z^c)^2]^{\frac{1}{2}} \quad \dots(5.17)$$

and X^c, Y^c, Z^c : provisional values of the receiver coordinates

The observation equation coefficients (partial derivatives) are derived by differentiating equation (5.13), as

$$\frac{\partial \ell_j}{\partial X} = \frac{X^c - X_j}{R_j^c}, \quad \frac{\partial \ell_j}{\partial Y} = \frac{Y^c - Y_j}{R_j^c}, \quad \frac{\partial \ell_j}{\partial Z} = \frac{Z^c - Z_j}{R_j^c} \quad \dots(5.18)$$

and

$$\frac{\partial \ell_j}{\partial \alpha_o} = c \quad (5.19)$$

A minimum of four observation equations similar to (5.15) can be solved by least squares to give the unknowns, ΔX , ΔY , ΔZ and $\Delta \alpha_o$, according to the principles outlined in Section 2.5.2 and in Appendix F. An error analysis can also be performed using the covariance matrix, σ_{xx} (equation (2.160)), where

$$\sigma_{xx} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix} \quad (5.20)$$

and σ_x^2 : variance of the receiver's X-coordinate

σ_{xt} : covariance of the X-coordinate and the clock bias, α_o

The covariance matrix can be used in order to select the satellites which provide the best geometry, before any observations are actually made. If σ_0^2 in equation (2.160) is assumed to be unity, the covariance matrix for any available satellite configuration can be evaluated as

$$\sigma_{xx} = N^{-1} \quad (5.21)$$

where N : normal matrix (see Section 2.5.2).

Several measures of precision have been defined (Miliken and Zoller, 1978), which can be derived from the elements of σ_{xx} . One such measure, the position dilution of precision (PDOP), is given by

$$\text{PDOP} = (\sigma_x^2 + \sigma_y^2 + \sigma_z^2)^{\frac{1}{2}} \quad (5.22)$$

If PDOP is multiplied by an estimate of the user-to-satellite range error it gives the estimated error in the user's position. So, the higher the PDOP for a given satellite configuration, the poorer the navigation solution will be. In order to select the satellites which provide the best geometry the receiver has to choose the configuration which ensures the minimum value of PDOP. The best achievable 4-satellite geometry is provided by 3 satellites at 120° intervals in the user's horizon plane (best 2-D intersection), with a 4th satellite directly overhead (best vertical position). Clearly, there is a conflict between the best geometry and the need to minimize refraction errors, because the low-elevation of the 3 satellites results in unacceptably high atmospheric effects. For this reason it is normal practice to avoid using satellites at below 5° elevations.

5.4.2 Pseudo-Range Measurements

The previous section described the basic real-time navigation technique using the GPS signals. The instantaneous accuracy is of the order of 15 m using the P code and 20-30 m using the C/A code (Parkinson and Gilbert, 1983). For geodetic positioning it is desirable to have the highest attainable accuracies. Such accuracies can be realized by observing the GPS signals over a number of hours, thus averaging out some of the errors that are inherent in the instantaneous solution. The satellite signals can be observed and processed in a number of ways, and the following three sections describe these high precision techniques.

This section will develop the mathematical models for pseudo-range measurements, since this is the main GPS observable. The pseudo-range model will depend on the available receiver system. Two possibilities will be considered here. The first is a receiver with a low-cost crystal oscillator which can observe 4 or more satellites simultaneously and will be denoted by R1. The second is a receiver system incorporating a highly stable atomic frequency standard which can only observe one satellite at a time and will be denoted by R2. The above receivers will also be considered in the other modes of observation of the GPS signals, in Sections 5.4.3 and 5.4.4.

The pseudo-range model for R1 is the same as for real-time navigation (see equation(5.13)) and is given by

$$\ell_j = R_j + c\alpha_o^{(i)} + E_{ion} + E_{trop} \quad (5.23)$$

A number, n_s , of sets of 4 or more simultaneous pseudo-ranges is observed over a few hours. The unknowns are one clock bias, $\alpha_o^{(i)}$,

for each set of simultaneous observations and the receiver antenna coordinates. There are a minimum of $4n_s$ observation equations for a total of

$$n_T = n_s + 3 \quad (5.24)$$

unknowns. These equations can be solved by least squares (see Section 2.5.2).

The solution can be made easier and the storage of large matrices can be avoided if a block-adjustment method of solution is adopted (Gough, 1978). In this method the normal equations are formed separately for each set of simultaneous observations. These can be expressed in matrix form as

$$\begin{bmatrix} n_{11} & n_{12} & n_{13} & n_{14} \\ n_{21} & n_{22} & n_{23} & n_{24} \\ n_{31} & n_{32} & n_{33} & n_{34} \\ n_{41} & n_{42} & n_{43} & n_{44} \end{bmatrix} \cdot \begin{bmatrix} \Delta\alpha_o^{(i)} \\ \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} \quad \dots(5.25)$$

or as

$$N_i x_i = d_i \quad (5.26)$$

where N_i : (4 x 4) normal matrix of the i^{th} set of simultaneous pseudo-ranges. This is the normal matrix used for the navigation solution.

x_i : (4 x 1) vector of unknowns for the i^{th} set of simultaneous pseudo-ranges

d_i : (4 x 1) right-hand-side vector

Since the clock offset is the only unknown which is unique in each set of normal equations, it can be eliminated by pre-multiplying both sides of equations (5.25) by

$$K_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{-n_{21}}{n_{11}} & 1 & 0 & 0 \\ \frac{-n_{31}}{n_{11}} & 0 & 1 & 0 \\ \frac{-n_{41}}{n_{11}} & 0 & 0 & 1 \end{bmatrix} \quad (5.27)$$

giving

$$\begin{bmatrix} n_{11} & n_{12} & n_{13} & n_{14} \\ 0 & & & \\ 0 & & M_i & \\ 0 & & & \end{bmatrix} \cdot \begin{bmatrix} \Delta\alpha_o^{(i)} \\ \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} = \begin{bmatrix} d_1 \\ f_i \end{bmatrix} \quad \dots \quad (5.28)$$

The reduced normals become

$$M_i x' = f_i \quad (5.29)$$

where M_i : (3x3) reduced normal matrix for i^{th} set of observations
 x' : (3x1) vector of unknown corrections to the provisional receiver coordinates

f_i : (3x1) reduced right-hand-side vector

Equations (5.29) are summed together (accumulated) for all sets of simultaneous observations, giving

$$\left[\sum_{i=1}^{n_s} M_i \right] x' = \sum_{i=1}^{n_s} f_i \quad (5.30)$$

or

$$Mx' = f \quad (5.31)$$

and the final solution for the receiver coordinates involves only the (3x3) matrix, M , of the accumulated normals. If the values of the clock biases are also needed, then the first equation in (5.28), for each set of observations, must be stored. When the solutions for the receiver coordinates (equation 5.31) have been obtained, these can be back-substituted to each of the stored equations in order to derive the clock biases. The maximum matrix size in the block-adjustment method is only (4x4), while an adjustment with all the normal equations will involve a matrix of dimensions $[(n_s + 3) \times (n_s + 3)]$. The computer storage requirements are therefore vastly reduced.

The pseudo-range model for R_2 has to take into account the behaviour of the receiver clock, because there are not enough observations to solve for a clock bias at every epoch. The need for a clock model implies that an atomic frequency standard is imperative for precise positioning. Equation (5.23) is modified as

$$\begin{aligned} \ell_j = R_j + c\alpha_0 + c\alpha_1(t - t_0) + c\alpha_2(t - t_0)^2 + E_{ion} + E_{trop} \\ \dots \quad (5.32) \end{aligned}$$

where α_0 : clock bias at epoch t_0
 α_1 : clock drift
 α_2 : aging term

In this technique a number of clock polynomials can be modelled. Since only one satellite can be observed at a time, the observational time span must be 4 times longer than for R1 in order to obtain a comparable amount of data.

Both techniques developed above assumed that the satellite clock predictions, Δt_{sj} (see equation (5.12)), for correcting the observed pseudo-ranges, ρ_j^0 , were correct. In the model for R1 there is no flexibility to allow for errors in the satellite clocks. In the model for R2 however, one can solve for a different clock polynomial for each satellite pass and avoid the need to correct the observed pseudo-ranges for the predicted satellite clock errors.

5.4.3 Phase (Doppler) Observations

A GPS receiver can make phase measurements of the carrier frequency by beating the satellite 'reconstructed' (L1 or L2) carrier (see Section 5.3.1) against a locally generated signal. At an observation epoch, τ , the phase of the receiver signal at site A is given by

$$\phi_A(\tau) = \phi_A(\tau_0) + 2\pi f_A(\tau - \tau_0) \quad (5.33)$$

where $\phi_A(\tau_0)$: phase (in radians) of locally generated signal at an initial (arbitrary) epoch, τ_0

$\Phi_A(\tau)$: phase of locally generated signal at epoch
of observation, τ (including whole cycles)

f_A : frequency of locally generated signal at A

At the same instant, τ , the phase of the received signal from the j^{th} satellite is

$$\Phi_j(\tau) = \Phi_j^A(\tau_0) + 2\pi f_j(\tau - \tau_0) - 2\pi f_j \left(\frac{R_{Aj}(\tau)}{c} \right) - \Phi_{\text{ion}}^A - \Phi_{\text{trop}}^A \quad \dots(5.34)$$

where $\Phi_j^A(\tau_0)$: arbitrary phase of carrier at the satellite (at τ_0)

$\Phi_j(\tau)$: received carrier phase from j^{th} satellite at epoch
of reception, τ

f_j : carrier frequency from the j^{th} satellite

$R_{Aj}(\tau)$: range between receiver at A, and satellite at the
time of transmission of the received signal

c : speed of electromagnetic waves in vacuo

$\Phi_{\text{ion}}^A, \Phi_{\text{trop}}^A$: phase arising from ionospheric and tropospheric
delays at A (see Section 5.4.5)

The receiver measures the phase difference between received and
locally generated signals as

$$\Phi_{Aj}(\tau) = \Phi_A(\tau) - \Phi_j(\tau) \quad (5.35)$$

which gives

$$\begin{aligned} \Phi_{Aj}(\tau) = \Phi_A(\tau_0) - \Phi_j^A(\tau_0) + 2\pi(f_A - f_j)(\tau - \tau_0) + 2\pi f_j \left(\frac{R_{Aj}(\tau)}{c} \right) \\ + \Phi_{\text{ion}}^A + \Phi_{\text{trop}}^A \end{aligned} \quad (5.36)$$

A receiver normally outputs Continuously Counter Doppler Counts (CCDC's), defined by

$$N_{Aj}(\tau) = \frac{\phi_{Aj}(\tau)}{2\pi} \quad (5.37)$$

The CCDC is the main geodetic observable and will be referred to as a phase measurement in this text. Equation (5.36) becomes

$$\begin{aligned} N_{Aj}(\tau) = & N_A(\tau_0) - N_j^A(\tau_0) + (f_A - f_j)(\tau - \tau_0) \\ & + f_j \left(\frac{R_{Aj}(\tau)}{c} \right) + N_{ion}^A + N_{trop}^A \quad \dots(5.38) \end{aligned}$$

where $N_{Aj}(\tau)$: CCDC or phase measurement

and all the cycle counts, N (eg N_{ion}^A), in equation (5.38) are defined by

$$N = \frac{\phi}{2\pi} \quad (5.39)$$

For precise measurements the observations are all made on the L1 frequency. The L2 observations are used in order to correct for first order ionospheric effects (see Section 5.4.5.1). The phase measurement of equation (5.38) can be converted into a biased range (similar to a pseudo-range) by multiplication with the wavelength of the carrier frequency, λ_j , where

$$\lambda_j = \frac{c}{f_j} \quad (5.40)$$

This gives

$$\begin{aligned} \rho_{Aj}(\tau) = & \frac{c}{f_j} (N_A(\tau_0) - N_j^A(\tau_0)) + c \left(\frac{f_A - f_j}{f_j} \right) (\tau - \tau_0) \\ & + R_{Aj}(\tau) + \lambda_j N_{ion}^A + \lambda_j N_{trop}^A \quad (5.41) \end{aligned}$$

where $\rho_{Aj}(\tau)$: biased range ($= \lambda_j N_{Aj}(\tau)$)

Equation (5.41) can be expressed in identical terms to equation (5.32) for a pseudo-range, as

$$\rho_{Aj}(\tau) = R_{Aj}(\tau) + c\alpha_0 + c\alpha_1(\tau - \tau_0) + E_{ion}^A + E_{trop}^A \quad (5.42)$$

where

$$\alpha_0 : \text{clock bias} \quad \left(= \frac{N_A(\tau_0) - N_j^A(\tau_0)}{f_j} \right)$$

$$\alpha_1 : \text{clock drift} \quad \left(= \frac{f_A - f_j}{f_j} \right)$$

An aging term, $c\alpha_2(\tau - \tau_0)^2$, can also be introduced in equation (5.42) if a frequency drift needs to be modelled.

Equation (5.42) is the appropriate mathematical model for phase measurements with the R2 receiver system (see Section 5.4.2). Its atomic clock can be adequately modelled by a constant frequency offset (clock drift) over quite long time intervals. Stabilities of $1 - 2$ parts in 10^{13} are common for cesium beam standards and these would result in errors of 30 - 60 cm after 3 hours, which is the approximate duration of a GPS satellite pass. R1 is a less appropriate system for this kind of observation, because the instability of the crystal oscillator will result in poor solutions even though this receiver can track all the visible satellites simultaneously.

The biased ranges of equation (5.42) are adjusted by least squares in order to solve for a number of clock biases, α_0 , a number of drifts, α_1 , and the receiver antenna phase centre WGS-72 coordinates. The observed biased ranges are given by

$$\rho_{Aj}^{O'}(\tau) = \frac{c}{f_j} N_{Aj}^O(\tau) \quad (5.43)$$

The observed phase measurement, $N_{Aj}^O(\tau)$, can take any initial value, since the initial phases of the receiver and satellite clocks, $N_A(\tau_0)$ and $N_j^A(\tau_0)$, are unknown. This means that the clock bias, α_0 , is different for each satellite pass and for every time that there is loss of lock on the signal phase. A number of bias unknowns, α_0 , have therefore to be introduced in order to account for all the instances of initial signal acquisition.

The observed biased ranges can, if required, be corrected for satellite clock errors using the information in the navigation message (see Appendix G.1). The corrected biased ranges are given by

$$\rho_{Aj}^O = \rho_{Aj}^{O'} + c\Delta t_{sj} \quad (5.44)$$

where Δt_{sj} : correction derived from navigation message (see equation (5.12))

If the corrected biased ranges are used, the drift term, α_1 , of equation (5.42) represents the clock drift (frequency offset) of the receiver oscillator with respect to GPS (predicted) time.

If the correction is not applied, the drift represents the frequency offset between receiver and satellite oscillators. In the latter case it is imperative to solve for one drift per satellite and in fact it is wise to introduce a new drift term per satellite pass in order to safeguard against changes in each satellite's clock.

In the formation of the observation equations (see Section 2.5.2) the computed biased ranges are derived from equation (5.42), using the computed values of the receiver coordinates and of the clock drift and bias terms. The observation equation coefficients

for the receiver coordinates and for the clock bias unknowns are given by equations (5.18) and (5.19) after replacing the pseudo-range, ℓ_j , for the biased range, $\rho_{Aj}(\tau)$. The observation equation coefficient for the drift term is

$$\frac{\partial \rho_{Aj}(\tau)}{\partial \alpha_1} = c (\tau - \tau_0) \quad (5.45)$$

For an adequate solution the receiver, R2, has to observe a number of satellite passes. Simultaneous observations on both L1 and L2 frequencies enable the computation of the first order ionospheric errors, E_{ion} (see Section 5.4.5.1), while measurements of pressure, temperature and relative humidity can be used to estimate the errors, E_{trop} , from tropospheric delays (see Section 5.4.5.2).

The only complication of the phase measurement (biased range) technique is the need to establish the GPS time of transmission of the satellite signals in order to compute the satellite coordinates from the orbital elements in the navigation message (see Appendix G). The GPS time can easily be derived by making use of the pseudo-range observations. If these are clocked at different local epochs from the phase measurements, then they have to be interpolated to the appropriate local times of reception. The GPS time of transmission of a signal that was received at a local epoch, τ , is given by

$$t_{tj} = \tau - \frac{\ell_{Aj}(\tau)}{c} \quad (5.46)$$

where $\ell_{Aj}(\tau)$: pseudo-range observation corrected for satellite clock errors and interpolated to local epoch of reception, τ .

The biased range approach has been adopted at Nottingham in order to process data from a geodetic receiver system which utilized an ST15010 GPS receiver (see Section 6.2.1). Chapter 6 gives the details of the software package and some of the results that were obtained.

The need to model clock bias parameters in the least squares adjustment can be eliminated if the phase measurements are converted to Doppler count observations. At an epoch, τ_1 , equation (5.38) gives

$$N_{Aj}(\tau_1) = N_A(\tau_0) - N_j^A(\tau_0) + (f_A - f_j)(\tau_1 - \tau_0) + f_j \left(\frac{R_{Aj}(\tau_1)}{c} \right) + N_{ion}^A(\tau_1) + N_{trop}^A(\tau_1) \quad (5.47)$$

At a second epoch, $\tau_2 (= \tau_1 + \Delta\tau)$, the phase measurement is

$$N_{Aj}(\tau_2) = N_A(\tau_0) - N_j^A(\tau_0) + (f_A - f_j)(\tau_2 - \tau_0) + f_j \left(\frac{R_{Aj}(\tau_2)}{c} \right) + N_{ion}^A(\tau_2) + N_{trop}^A(\tau_2) \quad (5.48)$$

The Doppler count is the difference between the phase measurements at τ_1 and τ_2 , given by subtracting (5.47) from (5.48) as

$$N_{Aj(1,2)} = (f_A - f_j)(\tau_2 - \tau_1) + f_j \left(\frac{R_{Aj}(\tau_2) - R_{Aj}(\tau_1)}{c} \right) + N_{ion}^A(1,2) + N_{trop}^A(1,2) \quad (5.49)$$

where $N_{Aj(1,2)}$: Doppler count between τ_1 and τ_2

$N_{ion}^A(1,2), N_{trop}^A(1,2)$: ionospheric and tropospheric errors

in the Doppler count.

Doppler count observations can be adjusted by least squares using the mathematical model of equation (5.49). They result in identical solutions to the biased range approach and have the advantage that there are fewer unknowns. The phase measurements, however, can easily be incorporated in interferometric methods of adjustment when simultaneous observations from two or more receivers are available. Interferometric techniques (see Section 5.4.4) eliminate common errors between the two or more sites and result in very precise relative coordinates. The full accuracy potential of GPS can only be realized when these techniques of adjustment are adopted.

5.4.4 Interferometric Techniques

The previous section described the biased range (phase measurement) technique for single point positioning. It was then explained that this technique is only suitable for receiver system R2 (see Section 5.4.2). Interferometric methods of adjustment provide relative coordinates between two sites (see fig 5.VI) and are based on two or more receivers making simultaneous observations to the same satellites (translocation). As will be seen in this section, these techniques are suitable for both the R1 and the R2 receiver systems.

The mathematical models will be developed first for R2. Equation (5.38) gives the phase measurement made by such a receiver system at site A. The corresponding simultaneous phase measurement from the same satellite at another site, B (see fig 5.VI), is

$$\begin{aligned}
 N_{Bj}(\tau) = N_B(\tau_0) - N_j^B(\tau_0) + (f_B - f_j)(\tau - \tau_0) + f_j \left(\frac{R_{Bj}(\tau)}{c} \right) \\
 + N_{ion}^B + N_{trop}^B
 \end{aligned}
 \tag{5.50}$$

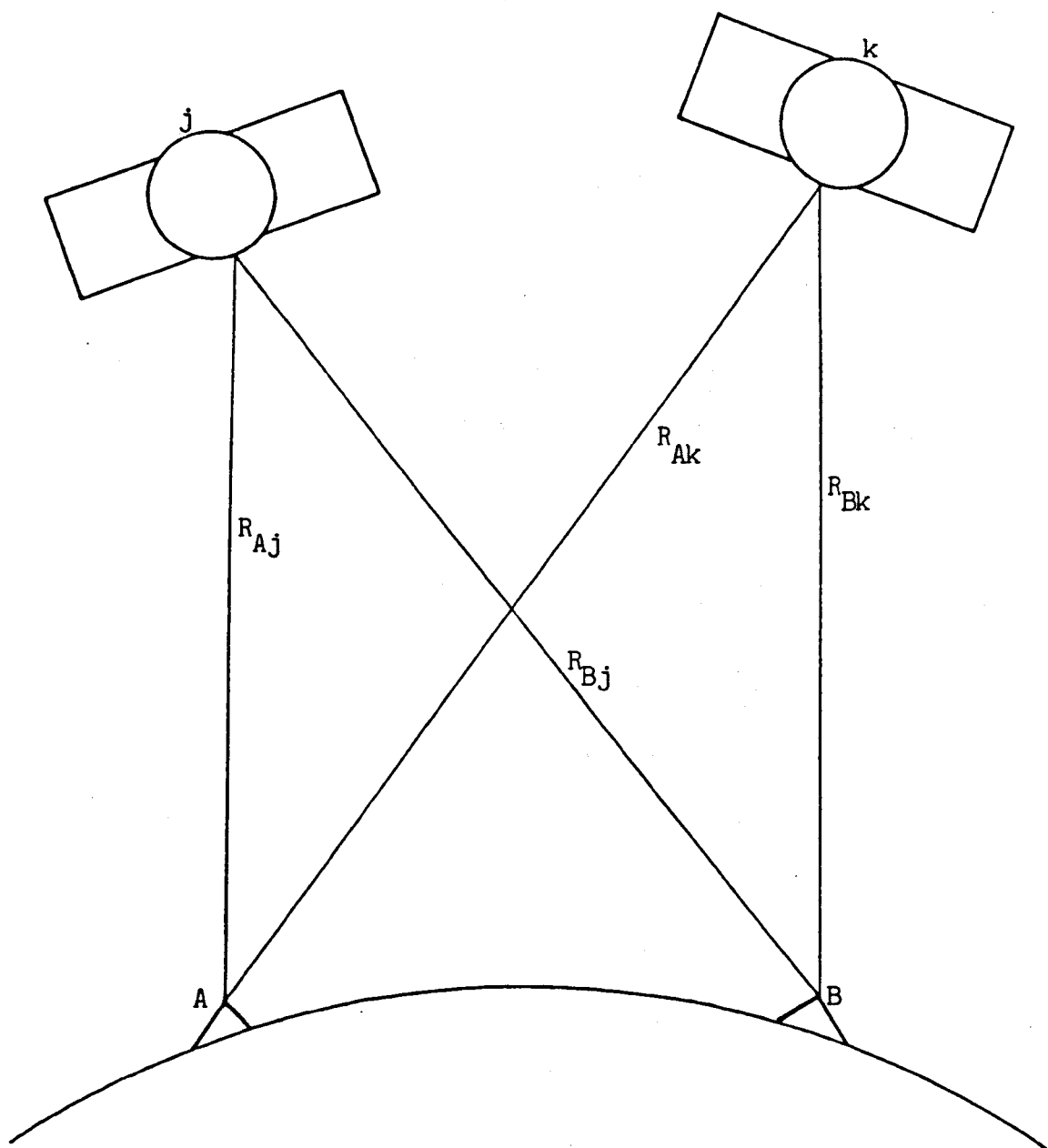


Fig 5.VI Satellite Interferometry

The main interferometric observable is the difference between the phase measurements at the two sites at any epoch, τ . This is called a single phase difference and is defined by

$$N_{ABj}(\tau) = N_{Bj}(\tau) - N_{Aj}(\tau) \quad (5.51)$$

which gives

$$N_{ABj}(\tau) = \alpha_{ABj} + (f_B - f_A)(\tau - \tau_0) + f_j \left(\frac{R_{Bj}(\tau) - R_{Aj}(\tau)}{c} \right) + N_{ion}^{AB} + N_{trop}^{AB} \quad (5.52)$$

Equation (5.52) is the mathematical model for a single phase difference measurement made by two receivers of the R2 type. The quantity

$$\alpha_{ABj} = (N_B(\tau_0) - N_A(\tau_0)) - (N_j^B(\tau_0) - N_j^A(\tau_0)) \quad (5.53)$$

represents an initial unknown bias. This has a different value for each satellite and also for when either receiver loses lock on phase (see Section 5.4.3). All instances of initial signal acquisition have therefore to be identified and bias unknowns introduced.

It can be seen in equation (5.52) that the phase difference observable is independent of all satellite clock errors. It therefore follows that there is no need for any corrections to be applied to the satellite clocks. In the least squares adjustment the phase difference observations from several satellite passes will be used in order to solve for a number of biases, α_{ABj} , a drift term (frequency offset), $f_B - f_A$, and the coordinate differences of the receivers at A and B.

The observation equation for a phase difference measurement, derived from equation (5.52), includes corrections to the approximate coordinates of both receivers. The interferometric approach is unable to derive absolute positions and thus it is necessary to fix the coordinates of one receiver by introducing position observation equations similar to equation (2.169). If absolute coordinates are required they can be obtained either from a pseudo-range or from a biased range adjustment (see Sections 5.4.2 and 5.4.3). These solutions can then be used as the approximate coordinates of the interferometric technique, holding one station fixed.

The drift unknown, $f_B - f_A$, represents the difference in frequency between the two receiver clocks. Its use is fully justified only in receivers with an atomic oscillator (e.g. R2), which can maintain a stable frequency. Any adjustments in either receiver clock within the time span of the observations have to be taken into account by the introduction of additional drift unknowns.

The observed value of the interferometric phase difference is given by

$$N_{ABj}^O(\tau) = N_{Bj}^O(\tau) - N_{Aj}^O(\tau) \quad (5.54)$$

where $N_{Aj}^O(\tau)$ and $N_{Bj}^O(\tau)$ are the observed phase measurements (CCDC's) of the two receivers at their respective local epochs, τ . If the biased range method is used first, it can be adapted to correct the observed phase measurements for ionospheric and tropospheric effects (see Section 5.4.5) and to identify the instances where loss of lock on signal phase has occurred. The

output of the two separate single-point adjustments can include the corrected phase observations, which are subsequently used as the input in the interferometric approach. In this approach the computed values of the phase difference observations can then be derived from equation (5.52), after excluding the terms N_{ion}^{AB} and N_{trop}^{AB} since these corrections have already been applied to the observed values.

The single phase difference approach is also suitable for receiver system R1, but a different mathematical model has to be developed. With this receiver system the phase measurement at site A is given by

$$N_{Aj}(\tau) = N_A(\tau_0) - N_j^A(\tau_0) + \epsilon_A(\tau) + (f_A - f_j)(\tau - \tau_0) + f_j \left(\frac{R_{Aj}(\tau)}{c} \right) + N_{ion}^A + N_{trop}^A \quad (5.55)$$

where $\epsilon_A(\tau)$: error in phase measurement (CCDC) arising from instabilities in the crystal oscillator at A

The simultaneous phase measurement at B is similarly given by

$$N_{Bj}(\tau) = N_B(\tau_0) - N_j^B(\tau_0) + \epsilon_B(\tau) + (f_B - f_j)(\tau - \tau_0) + f_j \left(\frac{R_{Bj}(\tau)}{c} \right) + N_{ion}^B + N_{trop}^B \quad (5.56)$$

The single phase difference at any epoch, τ , is

$$N_{ABj}(\tau) = \alpha_{ABj} + (\epsilon_B(\tau) - \epsilon_A(\tau)) + (f_B - f_A)(\tau - \tau_0) + f_j \left(\frac{R_{Bj}(\tau) - R_{Aj}(\tau)}{c} \right) + N_{ion}^{AB} + N_{trop}^{AB} \quad (5.57)$$

The receivers at the two sites make at least four simultaneous phase difference observations at any epoch, τ . The quantity

$$\beta_{AB}^{(i)}(\tau) = (\epsilon_B(\tau) - \epsilon_A(\tau)) + (f_B - f_A)(\tau - \tau_0) \quad (5.58)$$

is common to all simultaneously observed phase differences and the superscript, i , denotes the i^{th} set of simultaneous observations. The least squares adjustment uses a minimum number of $4n_s$ observations (where n_s is the total number of sets of simultaneously observed phase differences) to solve for a total of

$$n_T = n_b + n_s + 3 \quad (5.59)$$

unknowns. In the above equation n_b is the number of initial bias unknowns, α_{ABj} , which equals the total number of initial signal acquisitions, and n_s is the number of local oscillator error terms, $\beta_{AB}^{(i)}(\tau)$, which equals the total number of sets of simultaneous phase difference observations. The only other unknowns are the 3 coordinate differences of the two receivers.

From the basic single phase difference observable of equation (5.57), three further approaches will be developed here for receiver system R1. The first approach (Bossler et al, 1980 and Goad and Remondi, 1983) involves taking the differences between simultaneous phase difference observations from separate satellites. The double phase difference between satellites j and k is given by

$$N_{ABjk}(\tau) = N_{ABk}(\tau) - N_{ABj}(\tau) \quad (5.60)$$

giving

$$N_{ABjk}(\tau) = \alpha_{ABk} - \alpha_{ABj} + f_j \left(\frac{R_{Bk}(\tau) - R_{Ak}(\tau) - R_{Bj}(\tau) + R_{Aj}(\tau)}{c} \right) + N_{ion}^{ABjk} + N_{trop}^{ABjk} \quad (5.61)$$

It can be seen from equation (5.61) that the double phase difference, apart from the initial bias, is free from any satellite or receiver clock errors. The only unknowns in this case are the n_b bias unknowns and the 3 coordinate differences between the two sites. A further extension to the above technique is the triple phase difference method (Goad and Remondi, 1983) which takes the differences of the double phase differences from one observation epoch to the next. It can be easily seen that by taking triple phase differences the bias terms, $\alpha_{ABk} - \alpha_{ABj}$, are also eliminated and the only unknowns are the three coordinate differences.

It should be noted that the double phase and triple phase difference methods will, in theory, give identical results to the single phase difference approach. The only advantage of taking higher differences is the elimination of undesirable unknowns, resulting in an easier least squares adjustment with the minimum of manual interaction. However, it is often desirable to monitor the behaviour of the receiver clocks, in which case the oscillator error terms, $\delta_{AB}^{(i)}(\tau)$, are required. These can only be obtained from the single phase difference technique or from a new approach which is described below.

In the final technique the interferometric observable is the difference between the single phase differences from one epoch to the next, which will be referred to as the Doppler count difference. At τ_1 the single phase difference from the j^{th} satellite is

$$N_{ABj}(\tau_1) = \alpha_{ABj} + \beta_{AB}(\tau_1) + f_j \left(\frac{R_{Bj}(\tau_1) - R_{Aj}(\tau_1)}{c} \right) + N_{ion}^{AB}(\tau_1) + N_{trop}^{AB}(\tau_1) \quad (5.62)$$

At a second epoch, $\tau_2 (= \tau_1 + \Delta\tau)$,

$$N_{ABj}(\tau_2) = \alpha_{ABj} + \beta_{AB}(\tau_2) + f_j \left(\frac{R_{Bj}(\tau_2) - R_{Aj}(\tau_2)}{c} \right) + N_{ion}^{AB}(\tau_2) + N_{trop}^{AB}(\tau_2) \quad (5.63)$$

The Doppler count difference observable is defined by

$$N_{ABj(1,2)} = N_{ABj}(\tau_2) - N_{ABj}(\tau_1) \quad (5.64)$$

which gives

$$N_{ABj(1,2)} = \beta_{AB(1,2)} + f_j \left(\frac{R_{Bj}(\tau_2) - R_{Aj}(\tau_2) - R_{Bj}(\tau_1) + R_{Aj}(\tau_1)}{c} \right) + N_{ion}^{AB(1,2)} + N_{trop}^{AB(1,2)} \quad (5.65)$$

Equation (5.65) is free of the bias terms, α_{ABj} . Two receiver systems of type R1 can make at least 4 simultaneous Doppler count difference observations to different satellites between any two epochs, τ_1 and τ_2 . All simultaneously observed Doppler count differences have the oscillator error term, $\beta_{AB(1,2)}$, in common. This suggests the use of a block-adjustment method of solution, similar to that used for pseudo-ranges in Section 5.4.2. The normal equations are first formed for each set, i , of simultaneous observations. The error term, $\beta_{AB(1,2)}$, is unique to each set and can be eliminated from the normals, leaving a (3×3) matrix for the unknown corrections to the approximate values of the coordinate

differences. The reduced normals are summed (accumulated) for all the observation sets (see equation(5.30)) and the final solution gives the required coordinate differences. Back-substitution of the latter into the original normal equations can recover the individual oscillator error unknowns, $\delta_{AB(1,2)}$.

The interferometric techniques that were described above have been applied to a number of GPS receiver systems. A brief account of geoelectric receiver development is given in section 5.5. The trend is currently away from the R2 and towards the R1 receiver type. This is because atomic oscillators are expensive and difficult to maintain in the field and because the full GPS potential(strongest geometry)is only exploited when making simultaneous observations to all visible satellites.

5.4.5 Reduction of Observations

The received GPS satellite signals are subject to delays caused by the earth's ionosphere and troposphere and to relativistic frequency shifts. The ionosphere extends from an altitude of 40 Km to 400 Km and consists of ionized particles and electrons. The electron density varies along the signal path from one epoch to the next, introducing a variable delay to the satellite signals. This delay is frequency dependent and can be corrected by making simultaneous observations of the same satellite on two or more frequencies (see Section 5.4.5.1). The troposphere extends from the earth's surface to an altitude of about 40 Km. Its variable density causes changes in the refractive index which result in additional delays of the satellite signals. The models used for correcting tropospheric delays are based on measurements of surface pressure, temperature, and relative humidity at the time of observation (see

Section 5.4.5.2). The GPS corrections to relativistic frequency shifts are outlined in Section 5.4.5.3.

5.4.5.1 Ionospheric Corrections

A GPS signal arriving at a receiver experiences a delay because of the ionosphere. This delay is of the form

$$E_{\text{ion}} = \frac{A}{f^2} + \frac{B}{f^3} + \dots \quad (5.66)$$

where A, B, \dots : constants

E_{ion} : ionospheric range error

f : frequency of transmitted signal

The first order ionospheric term, A/f^2 , can be estimated by the use of two-frequency measurements.

The ionospheric correction will be derived first for pseudo-range observations (see Sections 5.4.1 and 5.4.2). The L1 pseudo-range is given by

$$\ell_{L1} = \ell_j + \frac{A}{f_{L1}^2} \quad (5.67)$$

where ℓ_j : pseudo-range corrected for ionospheric delays

f_{L1} : 1575.42 MHz

The simultaneous L2 observation is

$$\ell_{L2} = \ell_j + \frac{A}{f_{L2}^2} \quad (5.68)$$

where f_{L2} : 1227.60 MHz

Subtracting (5.67) from (5.68) gives

$$\ell_{L2} - \ell_{L1} = A \left(\frac{f_{L1}^2 - f_{L2}^2}{f_{L1}^2 \cdot f_{L2}^2} \right) \quad (5.69)$$

In the Global Positioning System the two frequencies are in the ratio

$$\frac{f_{L1}}{f_{L2}} = \frac{77}{60} \quad (5.70)$$

Substituting this in (5.69) gives the ionospheric range error at L1 as

$$\frac{A}{f_{L1}^2} = \frac{\ell_{L2} - \ell_{L1}}{\left(\frac{77}{60}\right)^2 - 1} = 1.54573 (\ell_{L2} - \ell_{L1}) \quad (5.71)$$

The corrected L1 pseudo-range observation is now given by

$$\ell_j = \ell_{L1} - 1.54573 (\ell_{L2} - \ell_{L1}) \quad (5.72)$$

In the least squares adjustment the pseudo-range observations could either be corrected as in equation (5.72) or alternatively the term A/f_{L1}^2 (equation (5.71)) could be added to the computed values as in equation (5.16).

The ionospheric range error is largely unpredictable but at the GPS frequencies it has a magnitude of the order of 10m. It should be noted that since L2 is a weaker signal than L1, it is often the case that the L2 pseudo-ranges are rather noisy. When the ionospheric correction is applied this effect is magnified by multiplication with the factor 1.54573 (see equation (5.72)), resulting in larger residuals than would be obtained if the uncorrected L1 pseudo-ranges were used.

In order to apply the ionospheric corrections to the phase measurements, the latter have to be first converted to Doppler counts (see Section 5.4.3). The observed L1 Doppler count (phase difference) between two epochs, τ_1 and τ_2 , is given by (see

equation (5.49))

$$N_{L1} = (f_A - f_j)_{L1} (\tau_2 - \tau_1) + f_{L1} \left(\frac{R_{Aj}(\tau_2) - R_{Aj}(\tau_1)}{c} \right) + \left(\frac{\Delta A}{f_{L1}^2} \right) \frac{f_{L1}}{c} \quad (5.73)$$

where $\Delta A = A(\tau_2) - A(\tau_1)$ (5.74)

and $A(\tau_1), A(\tau_2)$: ionospheric constants at τ_1 and τ_2
(see equation (5.66)).

The L2 Doppler count between the same two epochs is

$$N_{L2} = (f_A - f_j)_{L2} (\tau_2 - \tau_1) + f_{L2} \left(\frac{R_{Aj}(\tau_2) - R_{Aj}(\tau_1)}{c} \right) + \left(\frac{\Delta A}{f_{L2}^2} \right) \frac{f_{L2}}{c} \quad (5.75)$$

In most receivers the frequencies of the L1 and L2 local signals, f_{AL1} and f_{AL2} , are at the proportions given by equation (5.70).

This gives

$$N_{L2} = \frac{f_{L2}}{f_{L1}} (f_A - f_j)_{L1} (\tau_2 - \tau_1) + f_{L2} \left(\frac{R_{Aj}(\tau_2) - R_{Aj}(\tau_1)}{c} \right) + \left(\frac{\Delta A}{f_{L2}^2} \right) \frac{f_{L2}}{c} \quad (5.76)$$

Combining equations (5.73) and (5.76) gives

$$N_{L1} - \frac{f_{L1}}{f_{L2}} N_{L2} = \frac{\Delta A}{cf_{L1}} \left(\frac{f_{L2}^2 - f_{L1}^2}{f_{L2}^2} \right) \quad (5.77)$$

The corrected L1 Doppler count is given by

$$N_{Aj(1,2)} = N_{L1} - \frac{\Delta A}{cf_{L1}} \quad (5.78)$$

where the quantity $\frac{\Delta A}{cf_{L1}}$ is the same as $N_{ion(1,2)}^A$ in equation (5.49). By making use of equation (5.77), the corrected Doppler count becomes

$$N_{Aj(1,2)} = N_{L1} - (N_{L1} - \frac{f_{L1}}{f_{L2}} N_{L2}) \frac{f_{L2}^2}{f_{L2}^2 - f_{L1}^2} \quad (5.79)$$

and this reduces, using equation (5.70), to

$$N_{Aj(1,2)} = \frac{4620}{2329} \left(\frac{77}{60} N_{L1} - N_{L2} \right) \quad (5.80)$$

The corrected Doppler counts are either summed to give the corrected phase measurements (CCDC's) or they are adjusted either as Doppler counts (see Section 5.4.3) or as Doppler count differences (see section 5.4.4)

In some receivers the quantity

$$D = (f_A - f_j)(\tau_2 - \tau_1) \quad (5.81)$$

in equations (5.73) and (5.75), which corresponds to the zero-Doppler (zero range difference) case, is a constant for both frequencies. In this case D is first subtracted from both $L1$ and $L2$ Doppler counts to give

$$N'_{L1} = N_{L1} - D \quad (5.82)$$

$$N'_{L2} = N_{L2} - D$$

The new N'_{L1} and N'_{L2} are then used in place of N_{L1} and N_{L2} in equation (5.80) in order to obtain the corrected Doppler count.

5.4.5.2 Tropospheric Corrections

The tropospheric correction model that will be given here is known as the Hopfield simplified model (Gough, 1978) and it is mainly the result of work that was carried out by Helen Hopfield (1969, 1972 and 1977). In this model the tropospheric range error is given by

$$E_{\text{trop}} = \Delta s_d + \Delta s_w \quad (5.83)$$

where E_{trop} : tropospheric range error (metres) which has to be subtracted from the observed range

Δs_d : range error component arising from 'dry' troposphere

Δs_w : range error component arising from the water vapour content of the troposphere

The 'dry' component is computed from

$$\Delta s_d = \frac{77.6 \times 10^{-6} P h_{ds}}{5T \sin \left[(E^2 + 2.5^2)^{\frac{1}{2}} \right]} \quad (5.84)$$

where P : atmospheric pressure in mb

T : temperature in Kelvin

E : satellite elevation in degrees

h_{ds} : height in metres above the observing station (receiver) where the refractive index for the 'dry' component is exactly unity

The height, h_{ds} , has been determined by observations (Hopfield, 1977) and is given by

$$h_{ds} = 40110.0 + 148.81 (T - 273.16) \quad (5.85)$$

The satellite elevation is calculated from (Gough, 1978)

$$\sin E = \frac{1}{R} [\cos \lambda \cos \varphi (X_j - X) + \sin \lambda \sin \varphi (Y_j - Y) + \sin \varphi (Z_j - Z)] \quad \dots (5.86)$$

where λ, φ : geodetic longitude and latitude of receiver

X, Y, Z : geocentric cartesian coordinates of receiver

X_j, Y_j, Z_j : geocentric cartesian coordinates of satellite at the time of observation

R : distance between receiver and satellite

The 'wet' component of the tropospheric range error is

$$\Delta s_w = \frac{77.6 \times 4810 \times 10^{-6} e h_{ws}}{5 T^2 \sin[(E^2 + 1.5^2)^{\frac{1}{2}}]} \quad (5.87)$$

where e : partial water vapour pressure in mb

h_{ws} : height (metres) above observing station where the refractive index for the 'wet' component is exactly unity

The height, h_{ws} , varies between 8 Km and 13 Km. This variation is largely unpredictable and it is the major source of error in the tropospheric correction, even though the magnitude of the 'wet' component is only 10% of that of the 'dry' component. The value of h_{ws} is normally taken as

$$h_{ws} = 11000 \text{ m} \quad (5.88)$$

The partial water vapour pressure, e , is related to relative humidity or to wet-and-dry bulb temperatures by simple expressions (Sykes, 1979).

The maximum value of the 'wet' component, Δs_w , is at the 2 m level (above 7° elevation) with a percentage error (uncertainty) in the region of 30-40%. The maximum value of the 'dry' component, Δs_d , on the other hand, is around 20 m but the percentage error is only of the order of 1% (Black and Eisner, 1982).

The final value of the tropospheric range error, E_{trop} , is the sum of the two components, Δs_d and Δs_w . It is subtracted from the observed pseudo-range or biased range measurements to give the corrected observations. Alternatively it can be added to the computed values, as in equations (5.16) and (5.42). An observed phase measurement (see equation 5.38) is corrected for tropospheric delays by subtracting from it the quantity

$$N_{trop}^A = \frac{f_j}{c} E_{trop} \quad (5.89)$$

A Doppler count (see equation (5.49)) is corrected by subtracting from the observed count the quantity

$$N_{trop(1,2)}^A = N_{trop}^A(\tau_2) - N_{trop}^A(\tau_1) \quad (5.90)$$

where $N_{trop}^A(\tau_1)$, $N_{trop}^A(\tau_2)$: tropospheric corrections to the phase measurements at the start (τ_1) and at the end (τ_2) of the Doppler count, from equation (5.89)

5.4.5.3 Relativity

The GPS satellite clocks have to be corrected for drifts (frequency shifts) which arise because the satellites are at different gravitational potentials from the user (General Relativity), and because they travel at different velocities (Special Relativity). The variations in these frequency shifts with time, as they appear to the user, consist of a secular (constant) term and of a periodic term. The secular term is largely compensated by purposely setting the satellite clocks at a slightly lower frequency than the nominal 10.23MHz. In fact the satellite clock frequency is set low by a factor of 4.45×10^{-10} . Any remaining relativistic clock drifts are implicitly included in the satellite clock correction polynomials (see Appendix G.1).

The only other relativistic term is a variation in the satellite clocks arising from the different orbits. This term, Δt_{re} , is defined in Appendix G.1 and it varies in magnitude according to the eccentricity of the orbit. Δt_{re} is zero for a perfectly circular orbit ($e = 0$). For an eccentricity of 0.01, which is near the upper limit for GPS satellites, Δt_{re} has an amplitude of 20 nanoseconds (5m).

5.5 GPS Geodetic User Equipment

GPS geodetic receiver development started early during Phase I of the GPS program. Various US Government agencies, as well as a number of industrial concerns, sponsored the development of receiver systems which can process the GPS observables according to the techniques described in Section 5.4. A significant part of the receiver development effort has been focused on systems which do not require knowledge of the GPS codes. These are intended for civil

applications, in the event that future access to the P code is limited to military users.

The first operational geodetic receiver was the NAVSTAR Geodetic Receiver System (NGRS). This was developed by the Naval Surface Weapons Center (NSWC) with Defence Mapping Agency (DMA) sponsorship. The first unit, NGRS-1, was operational by February 1979, and an improved system, NGRS-2, was developed in 1980 (Hermann, 1981). The NGRS conforms with the R2 receiver type as defined in Section 5.4.2. It is based on a Stanford Telecommunications Inc.(STI) GPS receiver, connected to a cesium frequency standard, and can track one satellite at a time on both L1 and L2 frequencies. More details on this receiver system are given in Section 6.2.1.

The NGRS and another similar system developed by Sheltech Canada (Lachapelle and Beck, 1982) were put through a series of exhaustive tests. Lachapelle et al (1982) reported single point positioning accuracies of the order of 5 - 10m, using pseudo-range and Doppler (phase) observations over 10-15 hour time spans (see also Chapter 6). Relative positioning tests resulted in sub-metre accuracies for a 28 Km baseline, based on ^{~ 17 hrs.} 1000 minutes of data (Anderle, 1982). A long baseline experiment between two sites at Mahe in the Seychelles and at Smithfield, Australia, resulted in accuracies of the order of 1 part per million (ppm) (Evans and Hermann, 1982). In this experiment the major sources of error were shown to be due to uncertainties in the satellite ephemeris. A serious limitation was that only two satellites were used in the observations, resulting in rather poor geometry that did not change from day to day.

Secondly, they cannot decode the satellite navigation message.

Satellite ephemeris and timing information has thus to be input from some other source.

The first interferometric receiver is called the MacrometerTM and it is manufactured by Macrometrics Inc. (Counselman and Steinbrecher, 1982). It uses a crystal oscillator and makes phase measurements (see Section 5.4.3) on the L1 frequency from up to 6 satellites simultaneously. It can perform absolute positioning to accuracies of 10m, or better if an atomic frequency standard is used (Bock et al, 1983). The Macrometer was however designed as an interferometric receiver and it is when using these techniques that optimum geodetic accuracy can be achieved. It conforms to the R1 (see Section 5.4.2) specification of a GPS receiver system and can utilize all the interferometric techniques of Section 5.4.4 that were developed for this type of instrument. Extensive tests of the Macrometer (Hothem and Fronczek, 1983, Bock et al, 1983, Goad and Remondi, 1983) have repeatedly demonstrated relative positioning accuracies of under 2 ppm over time spans of 2-3 hours. A two-frequency instrument that can perform the first order ionospheric correction (see Section 5.4.5.1) is currently under development for the US Air Force Geophysics Laboratory. This is expected to provide an order of magnitude improvement in the relative positioning accuracies.

The second GPS receiver system that can operate without the codes is known as SERIES, which is an acronym for Satellite Emission Range Inferred Earth Surveying (MacDoran et al, 1982). It was originally developed at the California Institute of Technology Jet Propulsion Laboratory under NASA Sponsorship (MacDoran, 1979) and

uses a 1.5m uni-directional (dish) antenna and a rubidium frequency standard. SERIES observes one satellite at a time on both L1 and L2 frequencies. It sequences all available satellites, spending two minutes on each one. The main SERIES observable is a 'pseudo-range' (biased range). The difference between simultaneously observed 'pseudo-ranges' at two sites provides the interferometric observable which is used in the least squares solution. The main drawbacks of SERIES are the size of the dish antenna and the fact that it can only observe one satellite at a time. Very little material has been published on its accuracy. Melbourne (1983a) reports an accuracy of 0.4ppm for a 171 Km baseline. He also reports the development of a new system, SERIES-X, which uses a small omni-directional antenna.

It should be emphasised that GPS geodetic receiver development is still in its infancy. GPS is not yet operational and the full benefits of continuous 4-satellite coverage and strong geometry will only be available when the 18-satellite constellation is deployed. Major improvements in the accuracy of the satellite ephemeris computation, which is presently a large source of errors, are also expected to be made. Bearing these facts in mind, it is still too soon to speculate on the limits to the attainable relative positioning accuracy of GPS.

CHAPTER SIX

GPS SOFTWARE DEVELOPMENT AND

ANALYSIS OF SAMPLE DATA

6.1 Introduction

The research work at Nottingham on the Global Positioning System has been carried out in order to:

- (i) write software to process both pseudo-range and phase (Doppler) observations
- (ii) test the software using real data
- (iii) assess the accuracy of absolute positioning with GPS
- (iv) establish the theoretical background behind the interferometric techniques of adjustment (see Section 5.4.4)
- (v) be in a position to modify the software to use the interferometric techniques (see Section 6.4.3)

GPS data have been provided for Nottingham by DMA (Macomber, 1983). They comprise pseudo-range and phase observations taken at Dahlgren, Virginia, between 9 and 19 January 1981. The receiver used was the NGRS-2, (see Section 5.5). Details of this receiver system and of the GPS data are given in Section 6.2.

For several reasons the data could not be used in their original form. A pre-processing program, DMAPROC, was thus written in order to put the observations in a format suitable for further processing. Section 6.3 describes the main techniques involved in DMAPROC and gives a brief outline of the software. The GPS least squares adjustment program, GPSPROG, is described in Section 6.4. This takes the pre-processed data from DMAPROC and performs the solution for the receiver antenna coordinates and for a number of clock unknowns. It can work either with pseudo-range or with

phase observations. The results obtained from the Nottingham software using the DMA data are presented in Section 6.5.

6.2 THE GPS DATA

6.2.1 THE NAVSTAR Geodetic Receiver System (NGRS)

The NGRS (Hermann, 1981) is a GPS geodetic receiver system designed to operate primarily in the phase measurement (Doppler) mode. It was developed by the NSWG with DMA sponsorship (see Section 5.5). The observations that were provided for Nottingham were obtained by the second version of this receiver system, NGRS-2.

The GPS receiver in the NGRS-2 is the STI5010, which is manufactured by Stanford Telecommunications Inc. It is a single-channel instrument, capable of L1 and L2 pseudo-range and Doppler observations from one GPS satellite at a time. NGRS-2 operates with a Hewlett-Packard 5061A cesium clock. The receiver operation is controlled by an Intel 80/10 microprocessor. According to commands entered by the operator, the microprocessor controls the duration of the tracking period for each satellite.

The NGRS-2 observations are recorded on cassettes and are subsequently transferred onto 9-track magnetic tapes. Two such tapes have been delivered to Nottingham and the data contained in them are described in the next section.

6.2.2 Data Specification

The NGRS-2 data at Nottingham were recorded on two 9-track magnetic tapes. The tapes consist of four separate files, each containing the observations from one GPS satellite. The satellites involved had SV identities 5, 6, 8, and 9. The data were observed between 9 and 19 January 1981 and the observation schedule for each satellite is shown in fig 6.I.

The NGRS-2 files are divided into a large number of one-minute data blocks. Each block consists of five records, the contents of which are as follows :

- (i) record 1 contains the satellite and receiver identities and a number of pre-processing indicators which give information on the data contained in the remaining records of the one-minute data block
- (ii) record 2 consists of 10 sets of L1 and L2 pseudo-ranges observed every 6 seconds (at the satellite). The observations are time-tagged according to the GPS time of transmission.
- (iii) record 3 consists of 10 sets of L1 and L2 carrier phase measurements (CCDC's) observed every 6 seconds. This data has not been used in any of the adjustments.
- (iv) record 4 contains one set of L1 and L2 Continuously Counted Integer Doppler Counts (CCIDC's or phase measurements) observed every 60 seconds. These are time-tagged according to the local time of reception and are explained in more detail below. Record 4 also contains coded meteorological data for the

SATELLITE ID								
5			6		8		9	
DATE	FROM	TO	FROM	TO	FROM	TO	FROM	TO
9/1/81							23:30	
10/1/81	16:00	20:00	13:40	16:00	11:30	13:30		02:40
11/1/81	19:10	20:50					23:30	
12/1/81	16:20	20:50	14:20	16:00			23:20	02:40
13/1/81	16:00	20:50	13:40	15:50	10:50	13:30		02:20
14/1/81							23:20	
15/1/81	16:00	20:20	14:20	15:50	10:50	13:10	23:20	2:40
16/1/81	15:50	19:30	13:40	15:50	11:00	13:10	23:20	2:20
17/1/81	15:50	20:20	13:30	15:40	10:50	13:00	23:20	2:00
18/1/81	15:50	20:00	13:30	15:20	10:30	13:00		2:00
19/1/81	15:50	17:30	13:30	15:10	10:30	12:30	00:40	1:30

NOTES: Times given are UTC to the nearest 10 minutes

Fig 6.I NGRS-2 Observation Schedule

tropospheric correction (see Section 6.4.5.2) and the elements of the broadcast satellite ephemeris (see Appendix G.2)

- (v) record 5 gives the nominal local epoch of reception of the 60-second phase measurement in record 4. It also contains the satellite clock correction coefficients (see Appendix G.1)

The 60-second phase measurement is given in integer cycle units (see equation (5.38)). The time in seconds from the nominal local epoch, τ_i , to the epoch of the phase measurement (positive zero crossing), τ_i^m , is given in record 4. The correct local epoch of the phase measurement is

$$\tau_i^m = \tau_i + \Delta\tau_i \quad (6.1)$$

where τ_i^m : precise local epoch of phase measurement
 τ_i : nominal local observation epoch which coincides with the integer minute of local time
 $\Delta\tau_i$: time in seconds to positive zero crossing

The above relationship is illustrated in fig 6.II. In the Nottingham software (see Section 6.3) the L1 and L2 60-second phase measurements are interpolated to the integer minutes of local time, τ_i . This will assist future modifications of the software to utilize interferometric techniques (see Section 6.4.3). Successive phase measurements are first converted to Doppler counts (see Section 5.4.3) by

$$N_{(i-1,i)}^m = N(\tau_i^m) - N(\tau_{i-1}^m) \quad (6.2)$$

where $N(\tau_{i-1}^m), N(\tau_i^m)$: successive phase measurements (CCIDC's)

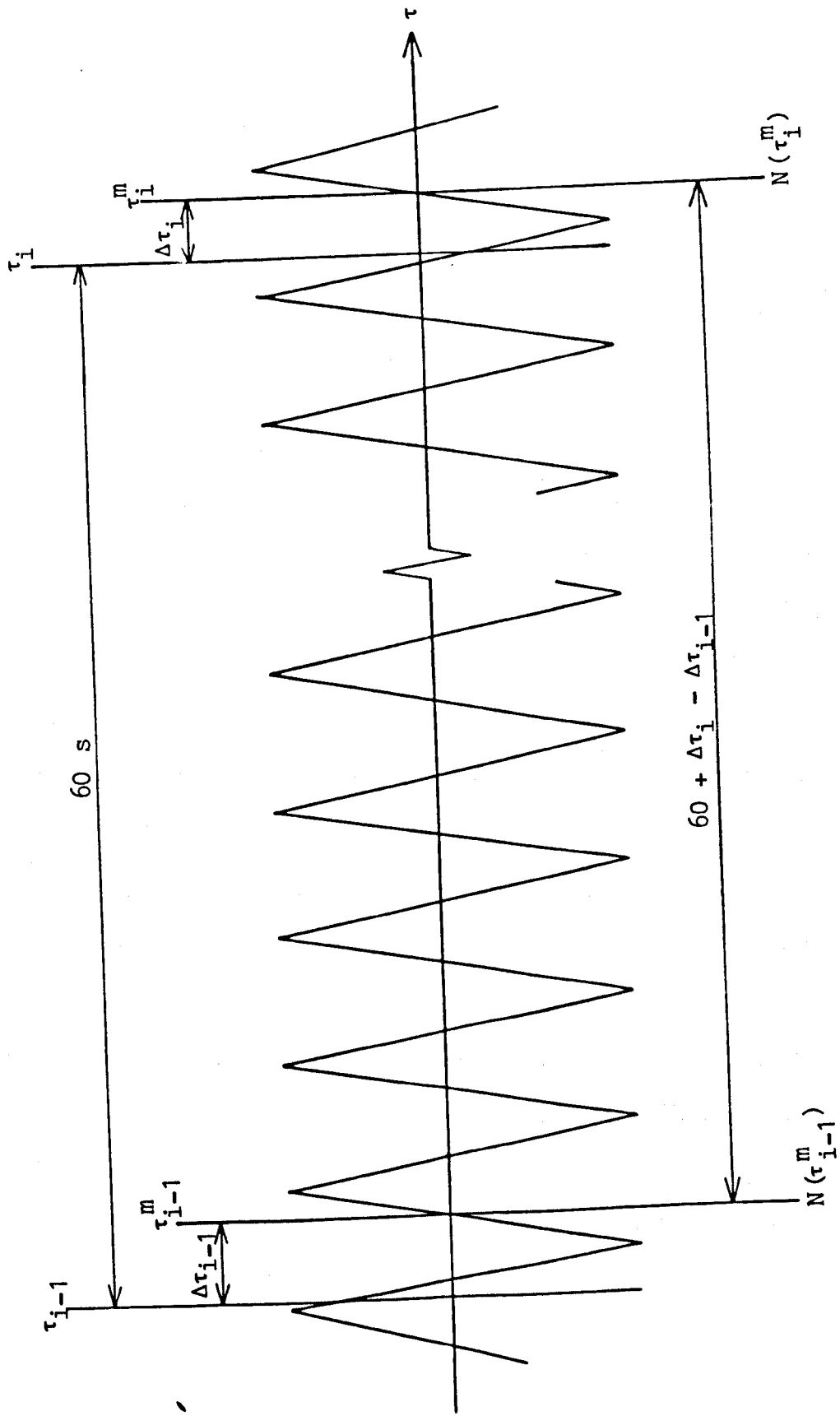


Fig 6.II NGRS-2 60 - Second Phase Measurement

The average Doppler frequency in the interval between the two phase measurements is (see fig 6.II)

$$f_{\text{Dop}} = \frac{N_{(i-1,i)}^m}{60 + \Delta\tau_i - \Delta\tau_{i-1}} \quad (6.3)$$

The modified Doppler count between two successive nominal (integer minute) local epochs is

$$N_{(i-1,i)} = N_{(i-1,i)}^m + (\Delta\tau_{i-1} - \Delta\tau_i) f_{\text{Dop}} \quad (6.4)$$

where the average Doppler frequency, f_{Dop} , is given by equation (6.3). Hermann (1981) estimated the maximum error in using the average frequency, f_{Dop} , as not exceeding 0.04 cm. The use of equation (6.4) for both the L1 and L2 Doppler counts results in these measurements being interpolated to common local epochs, thus enabling the estimation of the ionospheric correction (equation (5.80)). The final (corrected) counts can be added together to give the phase measurements at the integer minute intervals, corrected for ionospheric delays.

Associated with the two DMA data tapes is a third 9-track magnetic tape which contains the post-mission ephemerides of the four satellites. It comprises a set of cartesian WGS-72 satellite coordinates every 5 minutes. There are four files on this tape, each containing the ephemeris of one satellite. These ephemerides have been transferred onto random access files so that the satellite coordinates for any epoch can be computed as quickly and as easily as possible.

6.3 NGRS DATA PRE-PROCESSING SOFTWARE

6.3.1 Aims of Pre - Processing

The four NGRS-2 data files described in the previous section occupy a total of 16 megabytes of storage. Clearly, it is not practical to put such a large amount of data onto the computer's disk store, especially since a large amount of the information contained in the magnetic tapes is not required. In particular, the 6-second phase measurements in record 3 (see Section 6.2.2) have not yet been used. In addition, the elements of the broadcast satellite ephemeris are repeated in every one-minute block of data. These need to be retained whenever they are updated which is only once every 60 minutes. A pre-processing program was thus written in order to

- (i) vastly reduce the computer storage requirements by transferring only the necessary data to disk
- (ii) put the observations in a form suitable for the least squares adjustment program, GPSPROG (see Section 6.4)

A block diagram showing the main features of the GPS software at Nottingham is given in fig 6.III. The pre-processing program, DMAPROC, reads the NGRS-2 data files from the magnetic tapes and transfers the pre-processed observations to disk. The output of DMAPROC consists of a separate data file for each satellite. The final stage of pre-processing is to put the observations from all the DMAPROC files into a single file, in chronological order. This is performed by a short routine called DMASORT (see fig 6.III). This final file is the main data input for the adjustment program, GPSPROG.

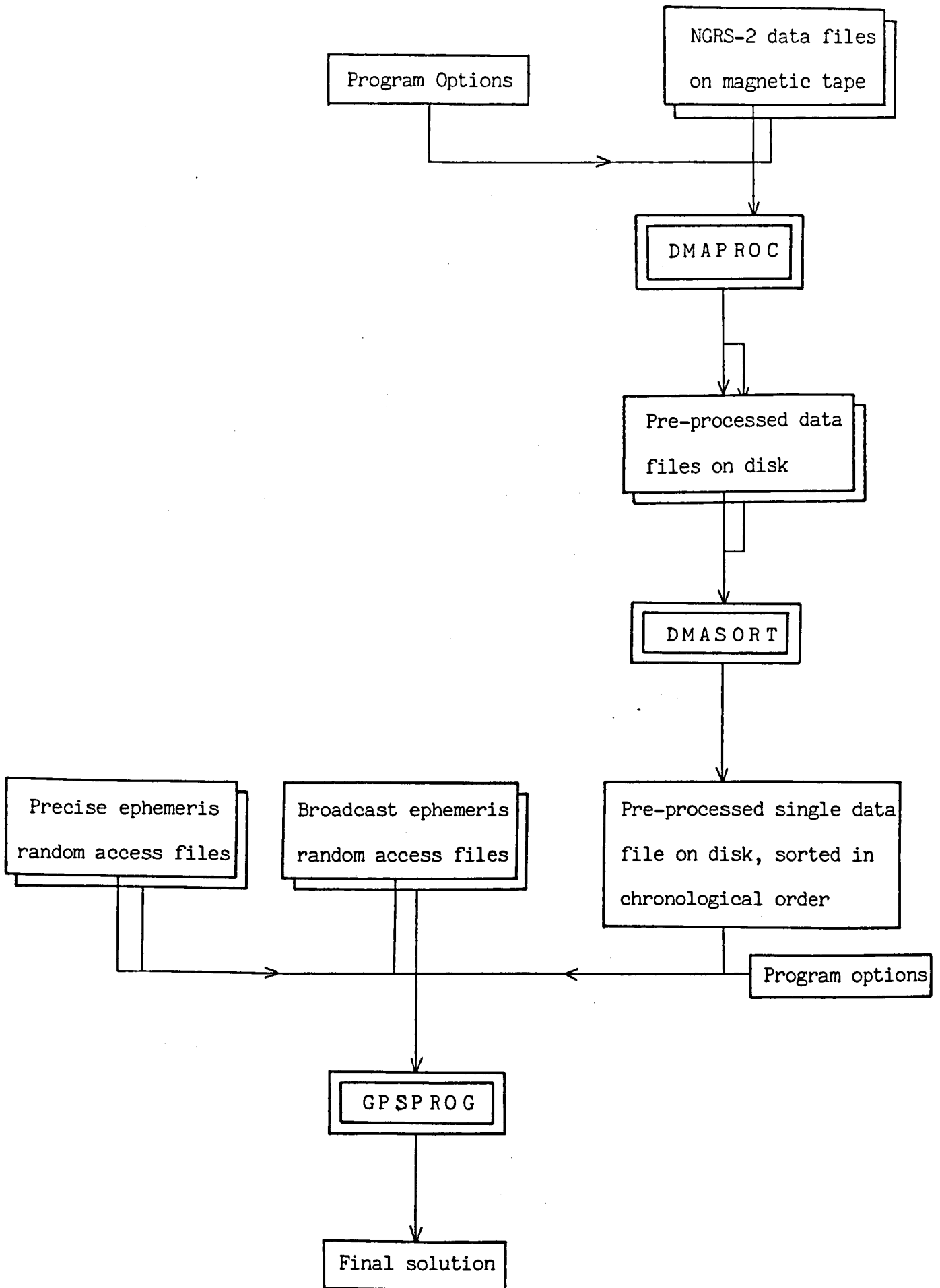


Fig 6.III Block Diagram of GPS Software

During the data pre-processing the broadcast ephemeris of each satellite is put into a separate random access file. By only storing the orbital elements whenever they are updated a significant reduction in the amount of data is achieved. Each pre-processed observation has an ephemeris record number associated with it. It is thus a simple matter to subsequently retrieve the appropriate satellite orbital elements from the random access file.

6.3.2 General Outline

The Nottingham GPS software (see fig 6.III) was written entirely in FORTRAN 77. The programs operate independently in two observation and adjustment modes. These are

- 1 : smoothed pseudo-range mode
- 2 : biased range (phase) mode

In mode 1 DMAPROC derives one smoothed pseudo-range measurement for every 4 minutes of pseudo-range data. The details of this mode and of the smoothing procedure are given in Section 6.3.2.1. The representation of up to 40 observations by a single measurement (normal point) results in a significant reduction in the storage requirements and facilitates the interpretation of the results (see also Section 3.1). In mode 2 DMAPROC converts the 60-second phase measurements defined in Section 6.2.2 to biased ranges. This mode is described in detail in Section 6.3.2.2.

The output of DMAPROC in either mode consists of one pre-processed data file per satellite. These data are sorted in chronological order into a single file by DMASORT (see fig 6.III). The final file is then used by GPSPROG to compute the WGS-72

coordinates of the receiver antenna, using the mathematical models developed in Sections 5.4.2 and 5.4.3.

The pre-processed biased range and smoothed pseudo-range files together occupy a storage of 1 megabyte. This is a 16-fold reduction in the amount of data after pre-processing. Further savings, probably by a factor of 2, can be achieved by additional refinements in DMAPROC.

6.3.2.1 Smoothed Pseudo-Range Mode

The NGRS-2 pseudo-ranges are observed every 6 seconds. This results in the accumulation of a large amount of data after tracking periods of a few hours. It was thus decided to use smoothed pseudo-ranges which are similar in concept to the normal points described in Section 3.1. In this approach a polynomial is fitted by least squares to a 4-minute span of pseudo-range observations. The smoothed pseudo-range is a point on the polynomial taken near the middle of the data span and it can represent up to 40 observations. In DMAPROC the order of this polynomial is specified by the user. It was found that a 5th order polynomial is best suited to the 4-minute data span and this results in rms residuals of the order of 1 metre.

Simultaneous smoothed pseudo-ranges are derived from both the L1 and L2 observations. An alternative approach would have been to apply the ionospheric corrections, using the two-frequency observations, before fitting the polynomial. This approach was discarded because it was found that a significant number of the L2 pseudo-ranges were rather noisy. It was thus preferable to leave the decision of whether or not to apply the ionospheric corrections until the solution stage (in GPSPROC). This has the advantage that solutions can be performed both with and without the ionospheric

corrections (see Section 6.5).

The operation of DMAPROC when it is in the smoothed pseudo-range mode is controlled by the input of a number of parameters. These are

- (i) the order of polynomial to be fitted to the pseudo-ranges
- (ii) the output mode. This varies the amount of line printer output that is available to the user.
- (iii) the dates and times between which data is to be processed
- (iv) the number of NGRS-2 files to be processed.

The NGRS-2 data files are processed in turn. Program operation starts with the input of successive one-minute blocks of data (see Section 6.2.2) from the first file on the magnetic tape. As soon as four consecutive one-minute blocks are read, least squares polynomials are fitted to both the L1 and the L2 pseudo-ranges. Each polynomial is of the form

$$l(t) = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 + \dots + a_n(t - t_0)^n + v$$

... (6.5)

where t_0 : arbitrary initial epoch

t : satellite time of transmission of the pseudo-range observation

$l(t)$: pseudo-range observation at time t

a_1, a_2, \dots, a_n : coefficients determined by least squares

v : least squares residual

n : order of polynomial

It was found that a 5th order polynomial best suits the 4-minute data span. A maximum of 40 observations are therefore used to determine the 6 unknown coefficients, a_0 , a_1 , a_2 , a_3 , a_4 , and a_5 . The smoothed pseudo-range is then derived from equation (6.5) by using a time argument, t , which is near the mid-point of the 4-minute data span. This epoch is the same for both the L1 and the L2 polynomials to enable the estimation of ionospheric delays. The program also calculates the rms residual, σ_R , of the fit of the observations to the polynomial. With a 5th order polynomial the value of σ_R is typically smaller than 1 metre. Noisy observations can be filtered by excluding from the solution all smoothed pseudo-ranges with a value of σ_R exceeding a certain limit.

While each one-minute block of data is being processed a check is maintained on the parameters of the broadcast satellite ephemeris. Whenever these are updated the new values are output in a random access file. A separate ephemeris file is assigned to each satellite and every record in this file contains all the necessary orbital and timing parameters (see Appendix G). These have a validity period of one hour. At the solution stage the appropriate satellite elements can be retrieved because DMAPROC outputs the ephemeris record number which corresponds to each smoothed pseudo-range.

The output of DMAPROC comprises one pre-processed data file for every NGRS-2 file. The pre-processed data files contain one smoothed pseudo-range observation per record and each record gives the following information:

- (i) satellite ID
- (ii) GPS week number (week 1 starts at midnight UTC on January 5 1980)

- (iii) the satellite time of transmission for the smoothed pseudo-range. This is given in seconds from the beginning of the week.
- (iv) the L1 smoothed pseudo-range
- (v) the simultaneous L2 smoothed pseudo-range
- (vi) the value of σ_R for the L1 pseudo-ranges
- (vii) the value of σ_R for the L2 pseudo-ranges
- (viii) the appropriate ephemeris file record number
- (ix) coded meteorological information

At the end of its run DMAPROC outputs a line printer listing summarizing the work it has carried out. For each NGRS-2 data file this listing gives:

- (i) the number of one-minute data blocks that were processed
- (ii) the number of smoothed pseudo-ranges
- (iii) the number of records in the resulting broadcast ephemeris file
- (iv) a printout of the ephemeris file

6.3.2.2 Biased Range (Phase) Mode

The biased range mode is the second mode of operation of DMAPROC. Here, the program converts the 60-second phase measurements (CCIDC's) defined in Section 6.2.2 to biased ranges (see Section 5.4.3). The 60-second phase measurements are far less noisy than the pseudo-ranges so it was decided to apply the ionospheric corrections at the pre-processing stage. In NGRS-2 the phase measurements are clocked at the local epochs of reception. However,

the satellite times of signal transmission are also required in order to enable the computation of the satellite coordinates from the elements of the ephemeris. These times are obtained by making use of the L1 pseudo-range observations. The pseudo-ranges are interpolated to the epochs of the phase measurements by making use of a least squares polynomial similar to the smoothing polynomial described in the previous section.

The operation of DMAPROC is controlled by the input of a number of parameters. These are identical to the ones listed in the previous section for the smoothed pseudo-range mode. The only additional information that is required by the program is a flag whose value determines whether or not the ionospheric corrections are to be applied. There are two available options here. The normal practice is for these corrections to be applied. Alternatively, the uncorrected L1 phase measurements can be used.

In the biased range approach (see Section 5.4.3) the observations at every initial signal acquisition can take any (arbitrary) values. These values are justified by the inclusion of bias unknowns at the solution stage (in GPSPROG). To easily identify the instances of initial lock on signal phase, DMAPROC sets the corresponding observed biased ranges to zero.

The program operation starts with the input of the first two one-minute blocks of data from the first NGRS-2 data file. The two consecutive L1 and L2 phase measurements (CCIDC's) are converted to Doppler counts using equation (6.2). These Doppler counts are then interpolated to the integer minutes of local time using equation (6.4) (see fig 6.II). The reasons why this is necessary

are twofold: firstly it results in the L1 and L2 Doppler counts being interpolated to identical epochs of reception, thus enabling the computation of the ionospheric corrections. Secondly, if simultaneously observed data from a second receiver are processed in the same way, they will result in the observations being interpolated to the same local epochs as those of the observations from the first receiver. It will therefore be a simple matter to further process the two data sets to obtain interferometric data types (see Section 5.4.4).

The NGRS-2 oscillator generates a local signal on both L1 and L2 frequencies which is offset by 28750 Hz from the GPS carriers. The 60-second zero-Doppler count (see equation (5.49)) for either the L1 or the L2 observations is thus

$$D = (f_A - f_j)(\tau_2 - \tau_1) = 28750 \times 60 = 1725000 \quad (6.6)$$

where f_A : local oscillator frequency (either L1 or L2)
 f_j : carrier frequency from j^{th} satellite (L1 or L2)
 τ_1, τ_2 : epochs of phase measurements (integer minutes)

In DMAPROC this zero-Doppler count of 1725000 cycles is subtracted from both L1 and L2 Doppler counts to give (see Section 5.4.5.1)

$$\begin{aligned} N'_{L1} &= N_{L1} - D \\ N'_{L2} &= N_{L2} - D \end{aligned} \quad (6.7)$$

where N_{L1}, N_{L2} : 60-second L1 and L2 observed Doppler counts, interpolated to integer minutes of local time
 (see Section 6.2.2)

The ionospheric correction is applied, if required, by substituting N'_{L1} and N'_{L2} for N_{L1} and N_{L2} in equation (5.80). The corrected Doppler count is thus given by

$$N_{(1,2)} = \frac{4620}{2329} \left(\frac{77}{60} N'_{L1} - N'_{L2} \right) \quad (6.8)$$

The program automatically sets the first observed biased range within a satellite pass, $\rho^0(\tau_1)$, to zero. The second observed biased range is given by

$$\rho^0(\tau_2) = \rho^0(\tau_1) + \frac{N_{(1,2)}}{f_{L1}} c \quad (6.9)$$

If the ionospheric correction is not required, the second observed biased range is obtained from

$$\rho^0(\tau_2) = \rho^0(\tau_1) + \frac{N_{L1}}{f_{L1}} c \quad (6.10)$$

After deriving the second biased range, DMAPROC reads the third one-minute block of data and repeats the process described above to obtain

$$\rho^0(\tau_3) = \rho^0(\tau_2) + \frac{N_{(2,3)}}{f_{L1}} c \quad (6.11)$$

Successive biased ranges are similarly derived until all the one-minute blocks of data have been processed.

For every 60-second Doppler count DMAPROC calculates

$$\begin{aligned} \Delta\rho_{L1} &= \frac{N'_{L1}}{f_{L1}} c \\ \Delta\rho_{L1} &= \frac{N'_{L2}}{f_{L2}} c \end{aligned} \quad (6.12)$$

where N'_{L1} , N'_{L2} : L1 and L2 Doppler counts corrected for receiver frequency offset using equations (6.7)

Equations (6.12) are used to identify instances of loss of lock on signal phase. It is assumed that the L1 phase measurement is affected differently from the L2 measurement when loss of lock occurs. The program then calculates the quantity

$$\Delta = | \Delta\rho_{L1} - \Delta\rho_{L2} | \quad (6.13)$$

If lock is maintained between two consecutive phase measurements then Δ is dependent only on ionospheric delays and its value is thus very small. So, when Δ exceeds a certain limit (arbitrarily set to 90m) it is assumed that loss of lock has occurred. In this case the observed biased range corresponding to the phase measurement of the latest one-minute block of data is reset to zero.

The observation epochs of the biased ranges are the local times of reception of the satellite signals. In DMAPROC the biased ranges are processed so that the observation epochs coincide with the integer minutes of local time (see Section 6.2.2). In order to enable the computation of the satellite coordinates, the satellite times of signal transmission are also required. These are obtained from the pseudo-range observations by making use of equation (5.46). For this equation to be valid, the times of reception of the pseudo-ranges must be the same as the observation epochs of the biased ranges. This condition is not satisfied with the NGRS-2 data. The pseudo-ranges must therefore be interpolated to the appropriate observation epochs and this is done by using a least squares polynomial fit of the form

$$\rho_{L1}(\tau) = b_0 + b_1(\tau - \tau_0) + b_2(\tau - \tau_0)^2 + \dots + b_n(\tau - \tau_0)^n + v \quad \dots \quad (6.14)$$

where τ_0 : arbitrary initial epoch
 τ : local time of reception
 $\ell_{L1}(\tau)$: L1 pseudo-range observation received at time τ
 b_0, b_1, \dots, b_n : coefficients determined by least squares
 v : least squares residual
 n : order of polynomial

The data spans for the polynomials are the same 4-minute intervals that are used in the derivation of the smoothed pseudo-ranges. The order, n , is specified by the user and, as in Section 6.3.2.1, a 5th order polynomial has been chosen. The six coefficients, b_0, b_1, \dots, b_5 , are derived by least squares for each 4-minute data span (40 observations). These coefficients are then used to derive the required pseudo-ranges by entering the appropriate observation epochs, τ , in equation (6.14). Each polynomial derives a maximum of four pseudo-ranges (one every integer minute) within the time-span of the observations that were used to compute it.

The accuracy of the pseudo-ranges in equation (5.46) depends on the speed of the satellites and on the accuracy requirements for the satellite coordinates. The GPS satellites travel at a speed of approximately 3800 ms^{-1} . So for an accuracy of 1 cm in the satellite coordinates the pseudo-ranges need only be known to within $2 \mu\text{s}$ (600 m). The L1 pseudo-ranges are therefore more than adequate for this purpose and the L2 observations are ignored in the derivation of the polynomials (equation (6.14)).

The operation of DMAPROC in the biased range mode is very similar to the program's operation in the smoothed pseudo-range mode (see Section 6.3.2.1). Each NGRS-2 data file results in one pre-processed biased range data file and in one satellite ephemeris

random access file (see fig 6.III). Each record in the biased range data files contains:

- (i) the satellite ID
- (ii) the GPS week number
- (iii) the satellite time of transmission of the biased range.
This is given in seconds from the start of the week.
- (iv) the biased range observation
- (v) the corresponding L1 pseudo-range (interpolated to the epoch of reception of the biased range)
- (vi) the rms residual of the polynomial fit to the L1 pseudo-ranges (see Section 6.3.2.1)
- (vii) the satellite ephemeris record number for the biased range observation
- (viii) coded meteorological data

The local epoch of reception, τ , for each biased range observation is not output by DMAPROC. This can be easily computed from

$$\tau = t + \frac{\ell_{L1}(\tau)}{c} \quad (6.15)$$

where the satellite time of transmission, t , and the L1 pseudo-range, $\ell_{L1}(\tau)$, are given in the pre-processed data files (see above).

The output of DMAPROC includes a line printer listing which summarizes the work that was carried out by the program. For each NGRS-2 data file this listing gives:

- (i) the number of one-minute blocks of data that were processed
- (ii) the number of biased ranges after pre-processing
- (iii) the number of ephemeris records in the resulting

random access file

- (iv) a printout of the ephemeris.

6.4 GPS LEAST SQUARES ADJUSTMENT PROGRAM (GPSPROG)

6.4.1 Introduction and Data Input

The final stage of data pre-processing (see Section 6.3) is the arrangement of all pre-processed observations in chronological order into a single file. This task is performed by a short routine called DMASORT (see fig 6.III). The two final data files, one for smoothed pseudo-ranges and the other for biased ranges, contain the pre-processed observations from all available satellites in the time span between 9 and 19 January, 1981 (see fig 6.I). These data are the main input for the least squares adjustment program, GPSPROG.

GPSPROG makes few distinctions between smoothed pseudo-ranges and biased ranges since the mathematical models for both data types are identical (see Sections 5.4.2 and 5.4.3). In both cases the program solves for the NGRS-2 antenna coordinates and for a specified number of clock(range) bias and clock drift unknowns. There are generally more bias unknowns with biased range observations than with smoothed pseudo-ranges because with the former data type the program has to solve for one such unknown for every initial signal acquisition.

During its run GPSPROG has to use data from a number of files (see fig 6.III). These are

- (i) the pre-processed data file (either smoothed pseudo-ranges or biased ranges)
- (ii) the four broadcast ephemeris random access files
(one per satellite)

- (iii) the four post-mission ephemeris random access files
- (iv) a disk file created by the user

The last file controls the operation of GPSPROG and contains :

- (i) the data mode (either pseudo-range or biased range)
- (ii) the ephemeris mode. The program can use either the broadcast or the post-mission ephemeris.
- (iii) the program output mode. This varies the amount of line printer output available to the user.
- (iv) a flag which determines whether the tropospheric corrections are to be applied
- (v) a flag which determines whether the L1 smoothed pseudo-ranges are to be corrected for ionospheric delays. If the biased ranges are to be employed this flag is used to tell the program whether the observations have been corrected for ionospheric delays in DMAPROC (see Section 6.3.2.2).
- (vi) a flag which determines whether the observations are to be corrected for predicted satellite clock errors. These corrections are computed from the information in the broadcast ephemeris (see Appendix G.1).
- (vii) the maximum admissible value of the rms residual, σ_R , of the polynomial fit to the pseudo-ranges (see Section 6.3.2). Observations with a value of σ_R which exceeds this limit are rejected.
- (viii) the input mode of the provisional receiver coordinates. Cartesian or spheroidal coordinates can be chosen.
- (ix) the semi-major axis and reciprocal of flattening of the required spheroid (if geodetic coordinates are chosen)

- (x) the provisional coordinates of the receiver (cartesian or geodetic)
- (xi) the number of bias unknowns
- (xii) the number of clock drift unknowns
- (xiii) the GPS epochs (week number and time of week) after which the program is to start solving for each bias unknown
- (xiv) the GPS epochs after which the program is to start solving for each drift unknown
- (xv) the provisional values of the bias terms
- (xvi) the provisional values of the clock drifts
- (xvii) the start and end epochs which define the time span of the observations to be used in the solution.

6.4.2 Program Description

GPSPROG starts its run by reading the information contained in the user-created disk file (see previous section). The data in this file specify the observation mode and control the operation of the program. A loop is then initiated where each pre-processed observation (biased range or smoothed pseudo-range) is entered in turn. For each data record the following tasks are performed :

- (1) The program checks whether the observation is within the required time span. If it falls inside this time span then it is accepted for further processing. Otherwise the next record is entered and the check is repeated until either an observation within the specified time span is found or the end of the data file is reached.

- (2) When the program is in the smoothed pseudo-range mode and if the ionospheric correction is to be applied then the rms residuals of the polynomial fits to the L1 and L2 observations are compared with the limit specified by the user (see Section 6.3.2.1). If either value of σ_R exceeds this limit then the smoothed pseudo-range is rejected and the next data record is entered. If the ionospheric correction is to be ignored then the L1 smoothed pseudo-ranges are employed and the above check is only performed for the rms residual of the fit to the L1 observations.
- (3) When the program is in the biased range mode then the rms residual of the polynomial fit to the L1 pseudo-ranges is also compared with the user-specified maximum limit. In this case the limit can be significantly larger than in (2) since the pseudo-ranges need only be accurate to within 600 m (see Section 6.3.2.2).
- (4) If required, the ionospheric correction is applied to the observed L1 smoothed pseudo-ranges (equation 5.72).
- (5) The broadcast satellite ephemeris record for the current observation is entered into the program.
- (6) The satellite time of transmission is corrected to the GPS time frame using the parameters in the navigation message (see Appendix G.1).
- (7) If the broadcast ephemeris is specified then the cartesian satellite coordinates are derived from the orbital elements and from the GPS time of transmission

(see Appendix G.2). If the post-mission ephemeris is chosen then the satellite coordinates are interpolated from the precise ephemeris random access file (see Section 6.2.2) by means of an 8th order Everett interpolation scheme (see Appendix E.2).

- (8) The program calculates the computed value of the range between the provisional location of the receiver and the satellite at the observation epoch (equation (5.17)).
- (9) The observation epoch is compared with the starting epochs of the bias and drift unknowns. The program then determines the locations of the coefficients for these unknowns within the observation equation and it locates the provisional values entered by the user (see Section 6.4.1).
- (10) The computed value of the observation is evaluated from

$$l^C = R^C + c \alpha_{or}^C + c \alpha_{1s}^C (t - t_{os}) \quad (6.16)$$

for a smoothed pseudo-range, or from

$$\rho^C = R^C + c \alpha_{or}^C + c \alpha_{1s}^C (\tau - \tau_{os}) \quad (6.17)$$

for a biased range, where

l^C : computed value of smoothed pseudo-range

ρ^C : computed value of biased range

R^C : computed value of range between receiver and satellite (see (8) above)

c : speed of propagation of electromagnetic waves

α_{or}^c : provisional value of r^{th} bias

α_{1s}^c : provisional value of the s^{th} clock drift

t_{os}, τ_{os} : GPS and local starting epochs for s^{th} drift
unknown

t : GPS time of transmission of smoothed pseudo-range

τ : local time of reception of biased range

- (11) If the tropospheric correction is required, the meteorological data are decoded to give the temperature, pressure, and relative humidity at the observation epoch. The tropospheric range error, E_{trop} , is then evaluated from equation (5.83) (see Section 5.4.5.2) and added to the computed value of the observation.
- (12) If the observations are to be corrected for satellite clock errors, the quantity $c\Delta t_s$ (see Appendix G.1 and equations (5.12) and (5.44)) is subtracted from the computed value.
- (13) The program calculates the observation equation coefficients for the receiver coordinates (equations (5.18)), for the clock bias (equation (5.19)), and for the drift term. The coefficient for the clock drift is given by

$$\frac{\partial \ell}{\partial \alpha_{1s}} = c (t - t_{os}) \quad (6.18)$$

for a smoothed pseudo-range, and by

$$\frac{\partial p}{\partial \alpha_{1s}} = c (\tau - \tau_{os}) \quad (6.19)$$

for a biased range. The coefficients are then placed in the appropriate locations of the observation equation vector.

- (14) The right-hand side of the observation equation is evaluated by taking the difference between the observed and the computed values.
- (15) The contribution of the observation equation is added directly to the normal equations. The observation equation is also placed in temporary store to enable the computation of post-adjustment residuals.
- (16) If the end of a satellite pass has been reached, the accumulated normal equations are solved to give the current solution.
- (17) If the end of the data file or of the specified time span (see Section 6.4.1) has been reached, the final solution is performed. Otherwise the next data record is read and the above process, starting from (1), is repeated.

The program output consists of the accumulated solution after every satellite pass, and of the final solution when all the data have been processed. The a posteriori standard errors of all the adjusted quantities are also given.

In the solutions, the receiver coordinates and their standard errors are given both in cartesian and in geodetic (spheroidal) representations. The ellipsoidal parameters for the latter (see

Appendix B.3) are either supplied by the user or default values are adopted. The default values are those of the WGS-72 ellipsoid, given by

$$a = 6378135.0\text{m} \quad (6.20)$$

$$\frac{1}{f} = 298.26 \quad (6.21)$$

GPSPROG solves for the cartesian receiver coordinates and a posteriori standard errors. The coordinates are then converted to the spheroidal representation as described in Appendix B.3. The standard errors are rotated to the spheroidal representation by applying Gauss's theorems (see Section 2.5.2) on equation (B.22). These give:

$$\sigma_{GG} = S^T \sigma_{xx} S \quad (6.22)$$

where σ_{xx} : (3x3) covariance matrix for cartesian receiver coordinates

S : (3x3) matrix defined in equation (B.21)

Equation (6.22) gives the covariance matrix, σ_{GG} , for the geodetic (spheroidal) receiver coordinates. Its diagonal elements are the variances, in units of length, along the directions of latitude, longitude, and height.

GPSPROG also outputs, if required, the post-adjustment residuals of the observations. These are given in a separate line printer listing which includes:

- (i) the satellite ID
- (ii) the GPS week number
- (iii) the observation epoch. This is the GPS time of transmission for smoothed pseudo-ranges, or the local time of reception for biased ranges.

- (iv) the observed smoothed pseudo-range or biased range
- (v) the computed value of the observation
- (vi) the 'observed minus computed' value
- (vii) the least squares residual

6.4.3 Future Development

The next stage in the software development at Nottingham is the extension of the GPS programs to incorporate interferometric techniques of adjustment (see Section 5.4.4). GPSPROG has been written with these techniques in mind and it can be easily modified if interferometric data are available.

In the final version of the software, simultaneous observations from two sites will be pre-processed by DMAPROC in exactly the same way as described in Section 6.3. The single point solutions, in the biased range mode, will then be performed independently for each site by GPSPROG. These solutions can be used as the provisional coordinates for the interferometric scheme. However, the function of GPSPROG will not stop here. This program will also be used as the final stage of data pre-processing. After each adjustment, GPSPROG will output the observed biased ranges and the computed values of the true ranges. It will also correct the data for all propagation delays.

The interferometric program will use the output of the two runs of GPSPROG as its main data input. By taking the difference of the simultaneously observed biased ranges between the two sites, an interferometric data type, the biased range difference, will be obtained. Also, by taking the difference of the corresponding computed ranges between the two single point solutions, the computed value of this observable will be derived (see equation (6.24) below). With little

extra work the program can thus obtain the right-hand side of the interferometric observation equation.

The GPS satellite coordinates will be needed for the computation of the observation equation coefficients. These are interpolated in GPSPROG to the GPS time of transmission of each biased range observation. They can then be output for use in the interferometric program, thus simplifying its operation by avoiding the need to incorporate the ephemeris files and associated computational procedures (see Appendix G).

The mathematical model for a biased range difference is obtained by multiplying both sides of equation (5.52) by c/f_j , as

$$\rho_{ABj}(\tau) = c\alpha_{op} + c\left(\frac{f_B - f_A}{f_j}\right)(\tau - \tau_{oq}) + R_{Bj}(\tau) - R_{Aj}(\tau) \quad (6.23)$$

This can be written as

$$\rho_{ABj}(\tau) = c\alpha_{op} + c\alpha_{1q}(\tau - \tau_{oq}) + R_{Bj}(\tau) - R_{Aj}(\tau) \quad (6.24)$$

where $\rho_{ABj}(\tau)$: biased range difference between stations A and B

α_{op} : p^{th} clock bias unknown

α_{1q} : q^{th} clock drift unknown

τ_{oq} : starting epoch of q^{th} drift unknown

$R_{aj}(\tau)$: range between receiver at A and satellite j, at the epoch of transmission of the biased range which was received at local epoch τ at A

$R_{Bj}(\tau)$: range between receiver at B and satellite j, at the epoch of transmission of the biased range which was received at local epoch τ at B

The computed values of the true ranges, $R_{Aj}^C(\tau)$ and $R_{Bj}^C(\tau)$, will already have been corrected for tropospheric delays in GPSPROC (see Section 6.4.2), while the biased range observations will have been corrected for ionospheric delays in DMAPROC (see Section 6.3.2.2). The interferometric program will use these data to solve for the receiver coordinates at the two sites, for a number of bias unknowns, and for a number of drift terms. The observation equation coefficients are derived by differentiating equation (6.24) with respect to each unknown (see Section 2.5.2). An interferometric adjustment only provides relative coordinates and thus the coordinates of one of the receivers will be held fixed by the introduction of appropriate 'observation equations' (see Section 2.5.3). The drift term in equation (6.24) has a very small dependence on satellite clock errors and it represents the frequency offset between the two receiver oscillators (see equation (6.23)). The biased ranges will not therefore need to be corrected for satellite clock errors in GPSPROC (see Section 6.4.2). A new bias unknown will have to be introduced in the adjustment every time either receiver initially acquires the satellite signals. Most of these instances will be identified in DMAPROC (see Section 6.3.2.2). Any cases of loss of lock on signal phase which are not detected in DMAPROC will be identified by examining the post-adjustment residuals from the single point solutions.

The GPS software at Nottingham has been developed for the NGRS, which conforms to the R2 receiver type (see Section 5.4.2). Programs have also been developed to process pseudo-range observations from receivers of the R1 type, such as the Geostar (see Section 5.5), according to the models discussed in Section 5.4.2. Additional software will be written to process interferometric data from R1 receiver systems, using the models discussed in Section 5.4.4, when such data are available.

6.5 ANALYSIS OF THE NGRS - 2 DATA

6.5.1 Introduction

The NGRS - 2 observation schedule is tabulated in fig 6.I. Two antenna sites were occupied during this data span. The WGS - 72 coordinates of site A were given by DMA (Macomber, 1983) as

$$\begin{aligned}\lambda_o &= 282^{\circ} 57' 21''.7110 \\ \varphi_o &= 38^{\circ} 20' 2''.5860 \\ h_o &= -37.266\text{m}\end{aligned}\quad (6.25)$$

Antenna site B was 2.000m south and 0.133m east of site A.

The pre-processed NGRS-2 observations at Nottingham were divided into 3 data sets which are summarized in fig 6.IV. Data Set 1 consists of observations from 12 satellite passes taken at site A between 23:30 UTC on day 9 and 20:40 UTC on day 14 (see fig 6.I). Data Set 2 also consists of 12 satellite passes at antenna site A. These were observed between 23:20 UTC on day 16 and 17:30 UTC on day 19. Data Set 3 consists of 8 satellite passes at antenna site B observed between 23:20 on day 14 and 19:30 on day 16. In each data set there are approximately 4 times more biased range than smoothed pseudo-range observations. This is because biased ranges are derived in DMAPROG at the rate of one every minute, while smoothed pseudo-ranges are derived at the rate of one every four minutes (see Section 6.3).

Single point solutions were performed, using the above data sets, in both the biased range and the smoothed pseudo-range modes of adjustment (see Section 6.4). The biased range solutions are

PASS	DATA SET 1 (SITE A)				DATA SET 2 (SITE A)				DATA SET 3 (SITE B)			
	SV ID	NO. OF BIASED RANGES	NO.* OF SMOOTHED PSEUDO-RANGES (L1 / L2)		SV ID	NO. OF BIASED RANGES	NO.* OF SMOOTHED PSEUDO-RANGES (L1 / L2)		SV ID	NO. OF BIASED RANGES	NO.* OF SMOOTHED PSEUDO-RANGES (L1 / L2)	
1	9	184	48/48		9	172	45/45		9	196	48/45	
2	8	114	32/23		8	140	37/36		8	56	20/5	
3	6	148	37/33		6	140	36/34		6	79	23/23	
4	5	227	59/59		5	279	71/71		5	243	59/58	
5	5	111	28/28		9	175	44/44		9	80	22/17	
6	9	188	49/49		8	140	38/36		8	116	30/25	
7	6	67	20/10		6	132	35/27		6	112	28/19	
8	5	239	61/61		5	279	69/69		5	215	51/51	
9	9	189	50/50		9	52	14/14					
10	8	104	25/0		8	163	39/39					
11	6	148	36/36		6	120	33/21					
12	5	280	70/70		5	140	35/35					

*Smoothed pseudo-ranges with a value of σ_R under 2m

Fig 6.IV Statistics of NGRS-2 Pre-Processed Observations

presented in Section 6.5.2 and the smoothed pseudo-range solutions in Section 6.5.3. Two different methods of modelling the receiver clock were attempted : the first method introduced a new clock drift (frequency offset) unknown for every satellite pass and did not correct the observations for the predicted satellite clock errors. Each drift parameter thus represented the frequency offset between receiver and satellite oscillators (see equation (5.41)). The second method applied the satellite clock corrections and only solved for one drift unknown for each data set. This unknown represented the local frequency offset from GPS (predicted) time.

Solutions were performed either with the broadcast or with a post - mission (precise) ephemeris (see Section 6.2.2). The latter was generated by NSWC using new (1983) software and the original Monitor Station data. A comparison between the two ephemerides is given in the next section.

Since it is most probable that only the L1 frequency will be available in the future for civilian use, solutions were attempted using the uncorrected L1 observations. The effects on the station coordinates of ignoring the ionospheric corrections are discussed in Section 6.5.4.

6.5.2 Biased Range Solutions

The least squares adjustment program, GPSPROG (see section 6.4), adds the contribution of each GPS observation directly on to the normal equations. GPSPROG performs the least squares solution at the end of each satellite pass. This solution contains the contributions of all previous observations and enables the user to examine the effect of each pass on the station coordinates.

The pass-by-pass biased range solutions for the three data sets, using the broadcast ephemeris, are tabulated in figs 6.V to 6.X. The station coordinates are given as corrections to the provisional values of equations (6.25). In the first three tables, figs 6.V to 6.VII, the adjustments included one drift unknown for each satellite pass. The observations were not corrected for satellite clock errors and thus the drift terms represented the frequency offsets between receiver and satellite clocks (see Sections 5.4.3 and 6.4.2). In the next three tables, figs 6.VIII to 6.X, just one drift unknown was incorporated in each adjustment. The observations were corrected for satellite clock errors and thus the drift represented the frequency offset between local and GPS clocks.

Fig 6.V shows the accumulated solutions after each pass for Data Set 1. The first satellite, ID 9, produced standard errors of the order of 30m and solutions which were within 300m of the DMA coordinates (equations (6.25)). The introduction of the biased range observations from the second pass improved the station coordinates and their standard errors by approximately one order of magnitude. The next significant improvement in the standard errors came after the 4th satellite pass, which produced a drop of 1.2m in σ_ϕ , accompanied by a 5m change in latitude. Subsequently, there was a very gradual drop in the standard errors up to pass 11. On the final pass there was a drop of 20 cm in σ_λ , accompanied by a change of 5.8m in the station longitude. Throughout the adjustment the longitude was the weakest component (i.e. produced the largest standard errors), and this weakness was reflected by the larger fluctuations in this coordinate than in latitude or height.

A feature of the GPS satellites is that they repeat the same ground track every 24 hours. This meant that successive passes of the same satellite in fig 6.V did not introduce any changes in the geometry and thus there was little improvement in the standard errors after pass 4. The rather large drop in σ_λ on the 12th pass can be attributed to the fact that this was 40 minutes longer than previous passes of satellite 5 (see figs 6.I and 6.IV) introducing some new geometry in the adjustment. When GPS is fully operational the 18 satellites will provide a much stronger geometrical configuration than was available at the time the NGRS-2 data were recorded. Significantly better solutions will then be possible, in much shorter time intervals.

An important result from fig 6.V is that the a posteriori standard errors with biased range data were over-optimistic. Drops of 10-20 cm in the standard errors reflected changes of 2-5 m in the station coordinates, implying that the standard errors were underestimated by at least a factor of 10. This situation can be rectified by using observations at less frequent intervals than the one-minute intervals of the biased ranges (see Section 6.5.3).

Fig 6.VI gives the pass-by-pass solutions for Data Set 2. These followed a very similar pattern to the solutions in fig 6.V. The only difference is that the longitude fluctuations were much smaller in fig 6.VI. This was because the first 8 passes were longer, on average, for Data Set 2 than for Data Set 1 (see fig 6.IV). The full range of geometry was thus sampled in the first 6-8 passes and there was little change in station coordinates or in standard errors after pass 8.

PASS	SV ID	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIIFT (from final solution) (x 10^{-12} sec/sec)
1	9	252.79	170.90	-65.90	32.75	23.41	9.18	0.42	1.4
2	8	10.24	-4.30	2.38	1.54	1.63	0.47	0.50	133.0
3	6	6.38	-2.01	0.09	1.40	1.70	0.48	0.64	34.6
4	5	8.02	3.25	-0.45	1.09	0.48	0.40	0.63	151.6
5	5	12.02	1.57	-0.55	0.94	0.43	0.39	0.62	153.5
6	9	14.02	0.52	-0.21	0.83	0.38	0.30	0.61	3.8
7	6	14.00	0.48	-0.16	0.81	0.37	0.29	0.60	32.8
8	5	16.38	-1.62	-0.04	0.70	0.32	0.30	0.62	153.3
9	9	16.83	-1.78	-0.13	0.65	0.31	0.28	0.61	3.6
10	8	17.27	-1.82	-0.32	0.62	0.30	0.27	0.60	132.1
11	6	17.32	-2.10	0.13	0.61	0.29	0.24	0.59	33.3
12	5	23.08	-3.78	0.02	0.41	0.27	0.26	0.64	153.4

NOTES : Observations not corrected for satellite clock errors
Number of Drift Unknowns : 12 (one per pass) Number of Biases : 23
Ephemeris : Broadcast

Fig 6.V Data Set 1 : Pass-by-Pass Biased Range Solutions

PASS	SV-ID	$\lambda - \lambda_0$ (m)	$\varphi - \varphi_0$ (m)	$h - h_0$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_0	DRIFT (from final solution) (x 10^{-12} sec/sec)
1	9	202.15	154.98	-54.02	70.69	50.00	18.97	0.53	1.0
2	8	12.88	55.15	-7.89	2.33	3.02	0.81	0.71	134.0
3	6	26.18	21.12	-1.97	2.54	2.67	0.78	0.86	29.1
4	5	27.70	-2.37	-3.79	0.82	0.60	0.78	0.94	156.2
5	9	27.14	-2.44	-2.98	0.68	0.54	0.58	0.85	1.6
6	8	24.39	-2.50	-0.11	0.62	0.54	0.50	0.86	131.9
7	6	24.33	-3.11	0.65	0.60	0.51	0.46	0.83	31.9
8	5	24.82	-5.26	0.91	0.50	0.44	0.47	0.84	155.5
9	9	24.83	-5.27	0.91	0.49	0.43	0.46	0.83	1.3
10	8	25.12	-5.11	0.43	0.44	0.42	0.37	0.81	133.1
11	6	25.13	-5.25	0.55	0.43	0.41	0.35	0.80	30.6
12	5	24.88	-4.97	0.57	0.41	0.38	0.34	0.78	152.8

NOTES: Observations not corrected for satellite clock errors

Number of Drift Unknowns : 12 (one per pass) Number of Biases : 23

Ephemeris : Broadcast

Fig 6.VI Data Set 2 : Pass-By-Pass Biased Range Solutions

PASS	SV ID	$\lambda - \lambda_O$ (m)	$\varphi - \varphi_O$ (m)	$h - h_O$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIFT (from final solution) (x 10^{-12} sec/sec)
1	9	350.78	209.62	-92.64	21.91	15.13	6.09	0.37	1.6
2	8	31.89	-11.39	-3.79	2.16	1.63	0.61	0.48	132.3
3	6	30.65	-12.70	-3.31	1.88	1.35	0.51	0.45	35.5
4	5	32.09	-3.51	-5.31	0.97	0.58	0.46	0.69	157.6
5	9	32.11	-3.62	-5.11	0.92	0.55	0.43	0.66	4.9
6	8	31.58	-3.50	-4.91	0.86	0.52	0.41	0.63	131.9
7	6	31.66	-3.71	-4.68	0.84	0.51	0.39	0.61	33.9
8	5	31.90	-5.58	-4.21	0.84	0.46	0.42	0.66	157.6

NOTES : Observations not corrected for satellite clock errors
Number of Drift Unknowns : 8 (one per pass) Number of Biases : 23
Ephemeris : Broadcast

Fig 6.VII Data Set 3 : Pass-By-Pass Biased Range Solutions

Fig 6.VII gives the solutions for Data Set 3 (antenna site B). Here, the satellite passes were rather shorter than for either of the other two data sets (see fig 6.IV). The standard errors after 8 passes did not therefore drop to the same level as in figs 6.V and 6.VI.

The drift values given in figs 6.V to 6.VII were the final solutions after all the observations had been incorporated. Since the observations had not been corrected for satellite clock errors, the drifts represented the relative frequency offsets between satellite and local oscillators (see equation (5.41)). It can be seen that the frequency offsets for successive passes of the same satellite were highly repeatable, with fluctuations of 2-3 psec/sec. This suggests that all oscillators were stable over the whole 10-day data span. Typical drift values for each satellite, from the satellite navigation message, were :

$$\begin{array}{rcll}
 & & \times 10^{-12} & \\
 \text{SV ID 5} & : & 152.5 & \text{psec/sec} \\
 \text{SV ID 6} & : & 31.5 & \text{psec/sec} \\
 \text{SV ID 8} & : & 136.0 & \text{psec/sec} \\
 \text{SV ID 9} & : & 3.0 & \text{psec/sec}
 \end{array} \tag{6.26}$$

The above values were relative to GPS time and were within 2-3 psec/sec of the NGRS-2 solutions in figs 6.V to 6.VII. It can therefore be inferred that the NGRS-2 clock drift relative to GPS time was negligible. This is substantiated by the solutions in figs 6.VIII to 6.X, where the observations were corrected for satellite clock errors. The value of the single drift unknown in each adjustment was very nearly zero.

Fig 6.VIII gives the pass-by-pass solutions for Data Set 1. On comparing these with the solutions for the same data set in fig 6.V, it can be seen that the standard errors were generally lower. The solutions were almost identical for the first satellite pass but after 4 passes the values of σ_λ and σ_φ were four times smaller than in fig 6.V. This was because there were fewer unknowns in the adjustment in fig 6.VIII, resulting in stronger geometry. After the 4th pass there was little change in the station coordinates and in the standard errors with the introduction of more observations. This was in contrast to the large longitude fluctuations in fig 6.V.

Since GPSPROG does not weight the observations, an a priori standard error of 1m is implied. The value of the variance of unit weight, σ_0 , is thus very nearly the value, in metres, of the rms residual of the observations. This value is, as expected, greater in fig 6.VIII than in fig 6.V, because of the smaller number of unknowns in the former adjustment. Sudden jumps in the value of σ_0 would be indicative of changes in the validity of the adopted mathematical models. For example, if there was a sudden increase in σ_0 after the inclusion of a new satellite pass in fig 6.VIII, this could be due to:

- (i) the satellite clock behaving differently from the GPS clock predictions
- (ii) a change in the frequency of the NGRS-2 clock, which should necessitate the inclusion of a new drift unknown
- (iii) errors in the satellite ephemeris

PASS	SV-ID	$\lambda - \lambda_O$ (m)	$\varphi - \varphi_O$ (m)	$h - h_O$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIIFT ($\times 10^{-12}$ sec/sec)
1	9	257.78	180.23	-73.75	32.39	23.16	9.08	0.42	45.0
2	8	21.44	4.49	-4.64	1.32	1.81	0.52	0.59	0.1
3	6	21.70	3.60	-5.29	0.78	0.13	0.42	0.65	-0.4
4	5	22.44	3.43	-6.94	0.25	0.12	0.39	0.77	-0.5
5	5	22.38	3.43	-6.27	0.24	0.11	0.33	0.73	-0.5
6	9	23.80	2.41	-5.66	0.30	0.14	0.37	0.97	0.0
7	6	23.78	2.39	-5.60	0.29	0.13	0.35	0.95	0.0
8	5	22.18	2.18	-5.27	0.24	0.14	0.36	0.98	0.1
9	9	22.37	2.11	-5.38	0.23	0.13	0.34	0.95	0.1
10	8	22.35	2.10	-5.19	0.23	0.13	0.33	0.94	0.1
11	6	21.55	1.37	-3.98	0.22	0.11	0.31	0.94	-0.1
12	5	19.51	0.92	-3.94	0.21	0.12	0.34	1.07	0.0

NOTES : Observations corrected for satellite clock errors

Number of Drift unknowns : 1

Number of Biases : 23

Ephemeris : Broadcast

Fig 6.VIII Data Set 1 : Pass-By-Pass Biased Range Solutions

PASS	SV ID	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIFT ($\times 10^{-12}$ sec/sec)
1	9	192.46	153.04	-57.19	69.52	49.17	18.65	0.52	42.9
2	8	31.47	35.72	-5.58	1.54	3.21	0.78	0.86	13.1
3	6	33.48	-2.17	-5.52	1.52	0.31	0.70	0.94	-3.7
4	5	13.17	-4.41	6.38	0.42	0.35	0.83	1.42	-1.9
5	9	13.86	-3.69	4.11	0.37	0.26	0.64	1.31	-1.8
6	8	13.23	-3.92	6.20	0.35	0.25	0.56	1.30	-2.1
7	6	13.49	-3.77	6.23	0.31	0.22	0.52	1.24	-2.0
8	5	13.82	-4.01	6.63	0.28	0.21	0.49	1.20	-2.1
9	9	13.86	-3.94	6.46	0.28	0.20	0.48	1.18	-2.1
10	8	14.16	-3.62	6.00	0.25	0.18	0.40	1.16	-2.1
11	6	14.03	-3.65	5.81	0.24	0.18	0.38	1.14	-2.1
12	5	14.83	-3.32	5.63	0.25	0.18	0.39	1.18	-2.0

NOTES : Observations corrected for satellite clock errors

Number of Drift Unknowns : 1 Number of Biases : 23

Ephemeris : Broadcast

Fig 6.IX Data Set 2 : Pass-By-Pass Biased Range Solutions

PASS	SV ID	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIFT ($\times 10^{-12}$ sec/sec)
1	9	337.76	205.86	-95.00	22.41	15.47	6.23	0.38	48.3
2	8	23.16	-15.95	-6.15	2.28	1.48	0.57	0.52	-6.4
3	6	31.26	4.31	-10.33	2.75	0.73	0.60	0.68	1.0
4	5	25.67	1.75	-8.80	0.40	0.28	0.50	0.80	0.6
5	9	25.34	1.47	-7.98	0.38	0.27	0.46	0.78	0.5
6	8	24.22	0.83	-5.66	0.38	0.27	0.44	0.81	0.2
7	6	23.90	0.58	-5.44	0.35	0.25	0.42	0.78	0.1
8	5	22.85	-0.29	-4.89	0.32	0.24	0.43	0.80	-0.1

NOTES : Observations corrected for satellite clock errors

Number of Drift Unknowns : 1 Number of Biases : 23

Ephemeris : Broadcast

Fig 6.X Data Set 3 : Pass-By-Pass Biased Range Solutions

Fig 6.IX gives the pass-by-pass solutions with only one drift unknown for Data Set 2. Again, the standard errors were generally smaller than those from the same data set in fig 6.VI, while the value of σ_0 was bigger. On comparing the repeatability between Data Sets 1 and 2, the final coordinates in figs 6.V and 6.VI agreed to within 2m, while the final coordinates in figs 6.VIII and 6.IX only agreed to within 9.5m, even though the standard errors in the latter two solutions were smaller.

Comparison of the solutions between figs 6.VII and 6.X, for Data Set 3, again shows the pattern of smaller standard errors with only one drift unknown. The final longitude in fig 6.X agreed more closely with the other solutions than with the longitude in fig 6.VII. This was probably due to the high value of σ_λ in fig 6.VII, because not enough passes had been observed to justify solving for one drift per pass.

From the above discussion it can be concluded that the solution scheme of allowing for the predicted satellite clock errors and solving for only one drift unknown is best for adjustments with just a few satellite passes as, for example, in Data Set 3. In all such schemes (figs 6.VIII to 6.X) there were relatively small changes in station coordinates after the 4th pass. When a large number of passes, with strong geometry, are available it is best to model the drifts and ignore the GPS clock corrections. This also safeguards the user from erroneous clock predictions made by the Control Segment.

The final solutions for station coordinates from figs 6.V to 6.X are tabulated again in fig 6.XI. This table also gives the final

solutions obtained using the precise ephemeris. If σ_0 is taken as a measure of the quality of the ephemeris used, it can be seen that there was little to choose between the two ephemerides. However, the variations between the different solutions were slightly smaller with the precise ephemeris, with a maximum range of 12.5m in each coordinate. The corresponding maximum range with the broadcast ephemeris was 17m.

Fig 6.XI shows a significant longitude bias between broadcast and precise ephemerides. The average coordinate differences between the two ephemerides for the six solutions, and their standard errors, were

$$\begin{aligned}\Delta\lambda_{B-P} &= -32''.039 \pm 0''.044 \\ \Delta\lambda_{B-P} &= 0''.197 \pm 0''.023 \\ \Delta\lambda_{B-P} &= -0.61\text{m} \pm 0.99\text{m}\end{aligned}\tag{6.27}$$

The above differences are in the sense 'precise minus broadcast'. The large longitude bias can be justified by a different origin of longitude having been adopted between broadcast and precise ephemerides. As explained in Section 2.5.4 one station's longitude has to be held fixed during any orbit determination process. The Naval Surface Weapons Center must therefore have changed the orientation of the GPS system between the two determinations, resulting in the 32'' longitude bias.

With either ephemeris it was impossible to detect the 2m latitude difference between antenna sites A and B. However, if more passes had been observed at site B it would have been possible to obtain more consistent solutions than those shown in fig 6.XI.

EPHEMERIS	DATA SET	NO. OF DRIFTS	$\lambda - \lambda_0$ (m)	$\varphi - \varphi_0$ (m)	$h - h_0$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_0
BROADCAST	1	12*	23.08	-3.78	0.02	0.41	0.27	0.26	0.64
	2	12*	24.88	-4.97	0.57	0.41	0.38	0.34	0.78
	3**	8*	31.90	-5.58	-4.21	0.84	0.46	0.42	0.66
	1	1	19.51	0.92	-3.94	0.21	0.12	0.34	1.07
	2	1	14.83	-3.32	5.63	0.25	0.18	0.39	1.18
	3**	1	22.85	-0.29	-4.89	0.32	0.24	0.43	0.80
	1	12*	-754.81	3.39	-1.58	0.44	0.29	0.28	0.68
	2	12*	-754.61	-2.29	4.59	0.38	0.35	0.32	0.72
	3**	8*	-746.57	1.21	-6.51	0.89	0.49	0.44	0.70
PRECISE	1	1	-758.98	7.04	-3.98	0.20	0.11	0.32	1.00
	2	1	-757.60	3.91	3.20	0.24	0.18	0.38	1.15
	3**	1	-752.76	6.29	-6.17	0.38	0.29	0.52	0.96

* Observations not corrected for satellite clock errors

** Solution for site B

FIG.XI Biased Range Final Solutions

When GPS is fully operational, and with the advent of receivers which can observe many satellites simultaneously, accuracies of 1-2m may be achieved using only a few hours' observations. The repeatabilities demonstrated by the NGRS-2 solutions, which can be taken as a measure of the accuracy, were at the 5-10m level.

6.5.3 Smoothed Pseudo-Range Solutions

The same types of solutions that were presented for biased ranges in the previous section were also attempted with the smoothed pseudo-range data. With this data type there were fewer bias unknowns since there was no requirement to keep track of signal phase. It was decided to incorporate one bias unknown per satellite pass.

Fig 6.XII gives the smoothed pseudo-range pass-by-pass solutions for Data Set 1 in an adjustment which modelled one drift per satellite pass. In this data set all the smoothed pseudo-range observations from pass 10 were rejected. This was because the L2 pseudo-ranges were too noisy and were filtered out of the adjustment (see fig 6.IV). The solutions can be compared with the corresponding biased range solutions in fig 6.V. It can be seen that the standard errors and the unit variance were always larger for the smoothed pseudo-range data. However, the solutions in fig 6.XII appear to have settled faster than the biased range solutions. After only 6 passes of the pseudo-range data there was no significant change in the station coordinates with the addition of more observations. In contrast, the longitude from fig 6.V changed by approximately 10m between pass 6 and pass 12. The final solutions from both data types agreed quite well in longitude and latitude but there was a rather large, 13m, difference in height. The final solutions for the

PASS	SV ID	$\lambda - \lambda_0$ (m)	$\varphi - \varphi_0$ (m)	$h - h_0$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o	DRIFT (from final solution ($\times 10^{-12}$ sec/sec)
1	9	-205.35	-160.42	78.95	242.46	168.89	67.84	1.98	6.5
2	8	16.31	-4.63	16.69	14.58	13.00	4.47	1.85	137.0
3	6	25.15	7.63	12.49	14.18	10.38	3.49	1.98	30.0
4	5	9.12	8.43	13.87	4.71	2.34	2.48	2.01	147.5
5	5	13.23	6.43	14.18	4.46	2.33	2.49	2.02	147.5
6	9	17.06	4.72	13.84	4.18	2.10	1.89	2.05	7.8
7	6	17.05	4.72	13.85	4.15	2.09	1.87	2.03	24.5
8	5	16.20	5.30	13.93	2.52	1.70	1.90	2.09	148.6
9	9	17.51	4.73	13.35	2.44	1.67	1.59	2.13	8.3
10	8	-	-	-	-	-	-	-	-
11	6	17.07	4.89	13.07	2.41	1.63	1.50	2.11	27.1
12	5	17.93	5.41	13.03	2.04	1.42	1.51	2.13	148.8

NOTES : Observations not corrected for satellite clock errors

Number of Drift Unknowns : 11 (one per pass) Number of Biases : 11

Ephemeris : Broadcast

Fig 6.XII Data Set 1 : Pass-By-Pass Smoothed Pseudo-Range Solutions

PASS	SV ID	$\lambda - \lambda_o$ (m)	$\psi - \psi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_ψ (m)	σ_h (m)	σ_o	DRIIFT ($\times 10^{-12}$ sec/sec)
1	9	-204.24	-153.88	72.18	241.94	168.53	67.69	1.98	-38.1
2	8	24.17	5.30	8.98	4.78	9.91	3.31	1.85	0.7
3	6	24.86	3.17	9.32	4.68	0.87	2.47	1.95	-0.2
4	5	23.37	3.40	9.51	1.13	0.78	2.32	2.19	-0.1
5	5	23.03	3.10	6.81	1.15	0.79	2.18	2.22	-0.2
6	9	23.60	2.39	7.34	1.06	0.72	1.75	2.24	0.1
7	6	23.59	2.37	7.17	1.05	0.71	1.73	2.22	0.1
8	5	24.42	2.83	6.97	0.99	0.70	1.73	2.30	0.1
9	9	25.14	2.43	6.60	0.95	0.66	1.53	2.35	0.4
10	8	-	-	-	-	-	-	-	-
11	6	22.80	0.68	8.89	0.83	0.56	1.43	2.39	-0.1
12	5	22.50	0.53	8.71	0.78	0.55	1.40	2.38	-0.1

NOTES : Observations corrected for satellite clock errors

Number of Drift Unknowns : 1

Number of Biases : 11

Ephemeris : Broadcast

Fig 6.XIII Data Set 1 : Pass-By-Pass Smoothed Pseudo-Range Solutions

drift terms were also very close and they were within 3-4 psec/sec of the values given in the satellite navigation message (see equations (2.26)).

Fig 6.XIII gives the pass-by-pass smoothed pseudo-range solutions for Data Set 1 in an adjustment which corrected the observations for satellite clock errors and solved for only one drift unknown. This can be compared with fig 6.VIII which contains the corresponding biased range solutions. In this scheme there was also good agreement in longitude and latitude and a 13m height difference between the two data types. The drift of the NGRS-2 clock relative to GPS time was again found to be negligible (see Section 6.5.2).

The final station coordinates with the smoothed pseudo-range data are tabulated in fig 6.XIV for all data sets. The same pattern as in fig 6.XI, that is, of smaller standard errors and larger rms residuals with the single drift unknown than with one drift per pass, was repeated. The standard errors were always larger for the pseudo-range, than for the corresponding biased range solutions in fig 6.XI. This was because the smoothed pseudo-ranges were less frequent (see Section 6.3) and noisier than the biased ranges. However, fig 6.XIV shows that the smoothed pseudo-range solutions were as consistent (repeatable) as the biased range ones. It can therefore be inferred that the standard errors in fig 6.XIV were more realistic than those from the biased range solutions. The former were still rather over-optimistic and they were probably under-estimated by a factor of 3.

Even though they were noisier, the smoothed pseudo-ranges resulted in solutions which were almost as good as the biased range solutions. The basic difference between the two data types

EPHEMERIS	DATA SET	NO. OF DRAFTS	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o
BROADCAST	1	11*	17.93	5.41	13.03	2.04	1.42	1.51	2.13
	2	12*	22.89	7.12	7.45	1.98	1.74	1.51	1.99
	3**	8*	29.53	-6.25	9.87	3.30	2.38	2.61	2.28
	1	1	22.50	0.53	8.71	0.78	0.55	1.40	2.38
	2	1	17.75	0.87	10.86	0.98	0.74	1.58	2.52
	3**	1	19.88	-2.05	7.55	1.55	1.17	2.58	2.41
	1	11*	-758.58	13.02	9.68	2.03	1.42	1.51	2.13
	2	12*	-760.16	10.24	13.87	1.85	1.63	1.41	1.87
	3**	8*	-747.55	-1.00	8.73	3.38	2.43	2.67	2.34
PRECISE	1	1	-755.68	7.42	6.41	0.79	0.55	1.41	2.40
	2	1	-756.01	6.51	7.96	0.82	0.62	1.34	2.13
	3**	1	-755.55	4.58	5.58	1.59	1.19	2.64	2.47

* Observations not corrected for satellite clock errors

** Solution for site B

Fig 6.XIV Smoothed Pseudo-Range Final Solutions

is that the pseudo-ranges are independent observations while the biased ranges are formed by summing successive Doppler-derived range differences. Any small errors in the latter data type can thus accumulate and result in unmodelled systematic effects. Indeed, the residuals from the biased range solutions exhibited systematic sinusoidal patterns which were not obtained with the smoothed pseudo-ranges.

The average coordinate differences between biased range and smoothed pseudo-range solutions are tabulated in fig 6.XVII. The solutions for longitude and latitude agreed quite well. However, there was a 10 m height bias between the two data types and the heights from the biased range solutions were in better agreement with the DMA value (equations (6.25)). The average coordinate differences in fig 6.XVII were almost identical for both ephemerides. This means that the differences between broadcast and precise ephemerides for smoothed pseudo-ranges were the same as for biased ranges and are given by equations (6.27).

6.5.4 Single-Frequency Solutions

The NGRS-2 observations in the adjustments described in the previous two sections had been corrected for ionospheric delays (see Section 5.4.5.1). In the future, it is projected that the L2 GPS frequency may not be available for civilian applications. This will necessitate the use of the L1 observations without ionospheric corrections. To study the effect of the ionospheric delays, the adjustments were repeated using the uncorrected L1 biased range and smoothed pseudo-range observations.

EPHEMERIS	DATA SET	NO. OF DRIPTS	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o (m)
BROADCAST	1	12*	15.71	-4.86	-15.71	0.40	0.26	0.25	0.61
	2	12*	22.31	-9.46	-16.49	0.35	0.33	0.30	0.67
	3**	8*	14.20	-7.06	-17.49	0.79	0.44	0.39	0.62
	1	1	17.43	-3.13	-19.59	0.23	0.14	0.38	1.20
	2	1	12.17	-6.92	-6.88	0.25	0.18	0.39	1.18
	3**	1	13.09	-10.33	-14.66	0.39	0.29	0.53	0.98
	1	12*	-762.18	2.32	-17.31	0.40	0.26	0.26	0.62
	2	12*	-757.18	-6.78	-12.48	0.35	0.32	0.29	0.66
	3**	8*	-763.61	-0.71	-19.93	0.82	0.45	0.41	0.64
PRECISE	1	1	-761.07	2.30	-19.63	0.21	0.12	0.35	1.10
	2	1	-760.26	0.30	-9.31	0.26	0.19	0.41	1.25
	3**	1	-762.73	-3.38	-15.87	0.48	0.35	0.65	1.21

* Observations not corrected for satellite clock errors

** Solution for site B

Fig 6.XV L1 Biased Range Solutions

EPHEMERIS	DATA SET	NO. OF DRIPTS	$\lambda - \lambda_o$ (m)	$\varphi - \varphi_o$ (m)	$h - h_o$ (m)	σ_λ (m)	σ_φ (m)	σ_h (m)	σ_o
BROADCAST	1	12*	28.24	-0.96	17.59	1.01	0.72	0.70	1.09
	2	12*	31.59	2.69	13.03	1.05	0.90	0.71	1.06
	3**	8*	30.09	-4.59	18.57	2.05	1.51	1.33	1.56
	1	1	24.18	3.87	12.30	0.42	0.30	0.69	1.46
	2	1	18.01	1.37	17.47	0.55	0.40	0.94	1.79
	3**	1	19.87	0.25	17.57	0.76	0.54	1.09	1.83
	1	12*	-748.84	6.46	14.99	1.02	0.73	0.71	1.11
	2	12*	-751.83	5.61	20.03	0.79	0.68	0.53	0.80
	3**	8*	-747.70	0.16	18.65	2.13	1.57	1.38	1.62
PRECISE	1	1	-754.22	10.44	10.05	0.43	0.30	0.70	1.48
	2	1	-753.52	8.73	12.25	0.44	0.32	0.75	1.44
	3**	1	-754.93	7.47	15.23	0.80	0.57	1.15	1.92

* Observations not corrected for satellite clock errors

** Solution for site B

Fig 6.XVI L1 Smoothed Pseudo-Range Solutions

The L1 biased range solutions are tabulated in fig 6.XV. On comparing these with the solutions which applied the ionospheric corrections, in fig 6.XI, it is seen that when solving for one drift per pass the rms residuals (σ_0) from the uncorrected observations were smaller than those from the corrected biased ranges. This was probably because the drift unknowns were absorbing the ionospheric effects and because the L1 observations were less noisy than L2. When solving for just one clock drift in each adjustment, the situation was reversed. In this case the rms residuals in fig 6.XV were the largest. This was probably because the single drift unknown could not adequately absorb the ionospheric effects from the uncorrected observations.

The coordinate solutions from the L1 biased ranges were different from those from the corrected observations. These differences had a maximum range of 18m. On average, the height component exhibited the greatest differences, with the L1 values being 16-17m smaller than the corresponding heights in fig 6.XI.

Fig 6.XVI gives the solutions computed using the L1 smoothed pseudo-ranges. The rms residuals were smaller than those from the adjustments which used the corrected smoothed pseudo-ranges, in fig 6.XIV. This was because the L2 pseudo-ranges were noisier than L1, thus contaminating the ionospheric corrections. The differences in station coordinates between figs 6.XIV and 6.XVI were smaller than those between the corresponding biased range solutions (figs 6.XI and 6.XV). The heights from the L1 pseudo-ranges were on average 6.5m larger than those from the solutions which applied the ionospheric corrections. In this case the change was in the opposite direction to that of the biased range solutions (see above).

EPHEMERIS	COORDINATE	WITH IONOSPHERIC CORRECTIONS		L1 FREQUENCY SOLUTIONS	
		AVERAGE VALUE	STANDARD ERROR	AVERAGE VALUE	STANDARD ERROR
BROADCAST	$\Delta\lambda$ (m)	- 1.10	1.36	9.52	1.61
	$\Delta\phi$ (m)	3.78	2.35	7.40	1.53
	Δh (m)	10.72	1.51	31.23	1.63
PRECISE	$\Delta\lambda$ (m)	- 1.37	1.37	9.33	1.74
	$\Delta\phi$ (m)	3.79	2.41	7.55	1.78
	Δh (m)	10.45	1.40	30.96	2.25

NOTES : Coordinate differences are in the sense 'smoothed pseudo-range minus biased range'

Fig 6.XVII Comparison Between Biased Range and Smoothed Pseudo-Range Solutions

Fig 6.XVII gives the average differences between the biased range and the smoothed pseudo-range coordinate solutions. When the ionospheric corrections had been applied, the only significant difference between the two data types was 10-11m in height. With the L1 observations there were significant differences in all components but again the height deviated most, by an average value of 31m. The other two components agreed to within 10m.

From the NGRS-2 data it appears that there was little loss of accuracy in the solutions for longitude and latitude when the ionospheric corrections had been ignored. However, there were significant differences in the height component between corrected and uncorrected observations. The effects of ionospheric delays should be further investigated, with an emphasis on isolating diurnal and seasonal variations and on establishing the maximum changes in station coordinates when using uncorrected observations.

CHAPTER SEVEN

CONCLUSIONS AND SUGGESTIONS FOR

FURTHER WORK

7.1 CONCLUSIONS

7.1.1 Conclusions on Orbit Determination

1. Orbit determination software has been developed at Nottingham and was successfully tested using observations of the LAGEOS satellite. Two 4-day sets of data were fully analysed.
2. Station coordinates were computed with precisions of 10-20 cm, with only limited amounts of data.
3. Agreement in station coordinates between solutions from two independent data sets and the SL5 solution was at the metre level when the BIH values of polar motion were used. The differences in station coordinates were reduced to under 40 cm after solving for polar motion.
4. The Nottingham polar motion solutions differed significantly (about 1 m) from those of the BIH. The former agreed to within 20 cm with the GEM-L2 polar motion values.
5. In separate orbit determinations, inaccuracies in the absolute values of earth rotation do not affect the solutions for station coordinates. It is important, however, to model changes in earth rotation (length of day) very accurately.
6. The currently used 5-day data spans for earth-rotation solutions are clearly inadequate, because UT1 - UTC cannot be assumed to remain constant over such long intervals.
7. The two LAGEOS data sets were used to derive daily changes in length of day (l.o.d.) with precisions of 0.0001 (5 cm). The Nottingham values agreed to within 0.0003 (15 cm) with the GEM-L2 solutions and to within 0.0001 with the BIH values.

8. Determination of the geocentric gravitational constant, GM , demonstrated a repeatability that was better than $0.01 \text{ km}^3 \text{ s}^{-2}$. This was also the level of agreement between the Nottingham, the SL5, the GEM-L2, and the MERIT values of the GM .
9. With the recently deployed 3rd generation laser tracking facilities it may be possible to achieve significant improvements in the attainable precisions of the orbit determination unknowns. The projected accuracy of station coordinates is 2-5 cm, enabling the monitoring of plate tectonic movements and of regional crustal deformations.

7.1.2 Conclusions on the Global Positioning System

1. This thesis has presented the theoretical basis for positioning with the Global Positioning System (GPS). Several models have been derived for the computation of both absolute and relative coordinates. For high accuracy work a special emphasis has been given to interferometric methods of adjustment.
2. A software package has been developed which can process pseudo-range and phase observations from the NGRS-2 receiver system and compute the absolute coordinates of the receiver antenna.
3. Various solutions were performed with NGRS-2 data. Two schemes of solution were attempted: the first corrected the observations for (predicted) satellite clock errors and solved for the clock drift of the local oscillator with respect to GPS time. The second scheme did not correct the observations but solved for one drift unknown per satellite pass.

It was seen that the first scheme was better for solutions with few satellite passes and rather poor geometry, while the second scheme was better for many passes and good geometry. In the latter, the drift values were in close agreement with those given in the satellite navigation message.

4. Absolute positioning with biased ranges (phase measurements) and either the broadcast or the post-mission (precise) ephemeris was repeatable to within 5-10m in adjustments which used data from 12 satellite passes. The a posteriori standard errors, which were of the order of 0.5m, were shown to be over-optimistic. There was a 32" longitude bias between the broadcast and precise ephemeris reference systems.
5. Absolute positioning with pseudo-range data was almost as good as with the biased ranges, even though the observations were noisier. The solutions for longitude and latitude were similar to the ones from the biased ranges but there was a 10m height bias between the two data types.
6. The omission of ionospheric corrections resulted in station coordinates which were within 20m of the solutions which included these corrections. The biggest differences were in the height component. The omission of ionospheric corrections affected the heights from the two data types differently; the biased range heights decreased while the pseudo-range heights increased. The average height difference between the two data types was 31m.

7. The NGRS-2 observations were made to just 4 GPS satellites. These repeated their ground tracks in successive passes. The geometry was therefore not significantly improved by the introduction of more than 4 satellite passes. Better solutions are to be expected when more satellites are available, thus providing a greater change in geometry. With the advent of geodetic receivers which can observe a number of satellites simultaneously, accuracies of 1-2m should be achieved within a span of just a few hours.

7.2 SUGGESTIONS FOR FURTHER WORK

7.2.1 Suggestions for Further Work on Orbit Determination

1. The Nottingham orbit determination software has been developed especially for LAGEOS. It should be extended to determine the orbits of other satellites (e.g. STARLETTE) by the addition of appropriate force model constituents, primarily air drag, depending on the satellite being tracked.
2. The optimum intervals for which to solve for C_R and C_a should be investigated. When integrating long arcs it may be necessary to introduce more than one value of each of these parameters.
3. The effect of the solar radiation cut-off (when the satellite crosses the earth's shadow) on the numerical integration routines (see Section 3.3.1) should be the subject of further study.
4. The analysis software (SOAP) should be modified in order to enable it to solve for a number of, say daily, polar motion and earth rotation values in one adjustment. This software should be used in order to establish the minimum time spans

for which the ERP's could be adequately determined.

Solutions should also be attempted which model drifts in earth rotation.

5. The data of the MERIT Campaign should be analysed to derive a good set of station coordinates and to compare the solutions for the ERP's with those of other computing centres.
6. The various improvements that will be brought about by using the 3rd generation laser tracking facilities should be investigated.
7. As more data is collected, an attempt should be made at improving the geopotential coefficients.

7.2.2 Suggestions for Further Work on GPS

1. The GPS software at Nottingham should be modified in order to incorporate interferometric techniques of adjustment, as described in Section 6.4.3.
2. Software should be written to process interferometric data from receivers of the R1 type (see Section 5.4.2), which can make simultaneous observations to a number of GPS satellites. These receivers (e.g. Geostar and Macrometer) can exploit the full potential of GPS to achieve the highest accuracies.
3. The effects of the ionosphere should be investigated further. For relative positioning, the indications are that sufficiently high accuracies are possible without the need for dual frequency measurements (Bock et al, 1983).

4. The validity of the tropospheric models should be investigated.
5. NAVSTAR GPS is still in its early stages. Further work will certainly be necessary in order to keep up with future developments as the system evolves.

APPENDIX A

ROTATION AND REFLECTION MATRICES

APPENDIX A

ROTATION AND REFLECTION MATRICES

If a rotation about the i^{th} axis of a right-handed coordinate system through an anti-clockwise angle θ , as viewed from the positive end of the axis towards the origin, is denoted by $R_i(\theta)$, then the following expressions define the three rotation matrices (Krakiwsky and Wells, 1971):

$$R_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

$$R_2(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R_3(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation matrices do not commute. The product of several rotations is performed from right to left. For example in

$$R_1(\alpha) R_2(\beta) R_3(\gamma)$$

the rotations are performed about the 3-axis of the original system, the 2-axis of the transformed system, and the 1-axis of the doubly

transformed system to yield the final system.

If a reflection of the i^{th} axis is denoted by P_i , then the following expressions define the three reflection matrices:

$$P_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Reflection matrices commute (e.g. $P_3 P_2 = P_2 P_3$). An odd number of reflections changes the handedness of the coordinate system.

APPENDIX B

COORDINATE REPRESENTATIONS

APPENDIX B

COORDINATE REPRESENTATIONS

B.1 Cartesian Representation

A cartesian representation of the position vector of a point, P, relative to an origin of coordinates, O (which could be the geocentre for earth-centred systems), is given by

$$\underline{P} = (X_p, Y_p, Z_p) \quad (\text{B.1})$$

where X_p, Y_p, Z_p : displacements from the origin, O, along 3 mutually orthogonal axes X, Y and Z (or 1, 2 and 3)

The coordinate axes can form either a right-handed or a left-handed system, as shown in fig. B.I. Most of the commonly used coordinate reference systems are right-handed, with the topocentric system being the most notable exception.

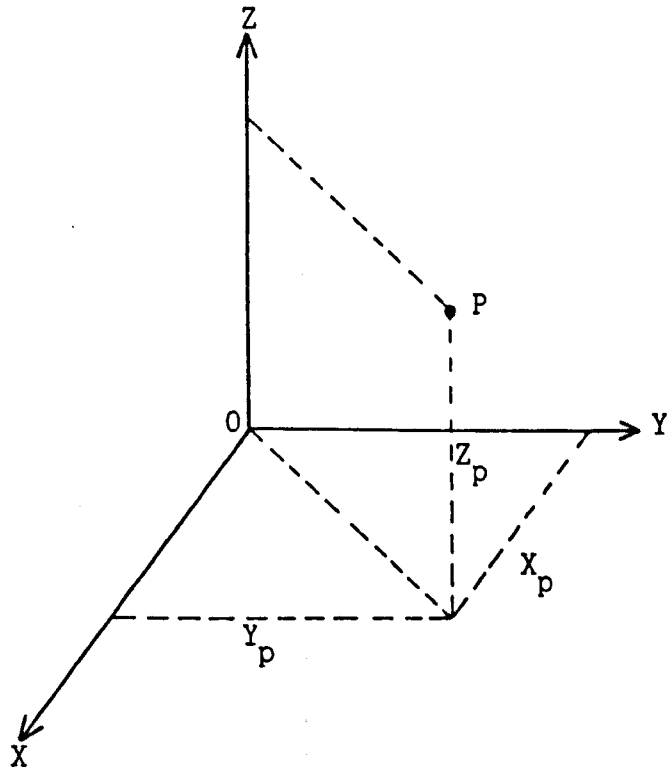
B.2 Spherical Polar Representation

The spherical polar representation of the position vector of a point, P, relative to an origin, O, is illustrated in fig.B.II and is given by

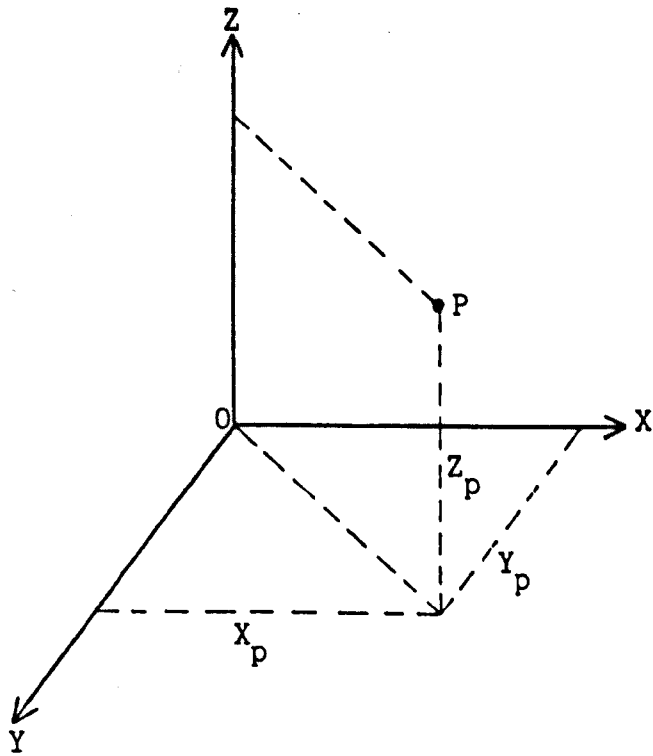
$$\underline{P} = (R, \Lambda, \phi) \quad (\text{B.2})$$

where R : distance from O to P

Λ : angle measured anti-clockwise from the X-axis to OA,
the projection of OP in the X - Y plane (longitude)



Right-Handed System



Left-Handed System

Fig B.I Cartesian Coordinate Systems

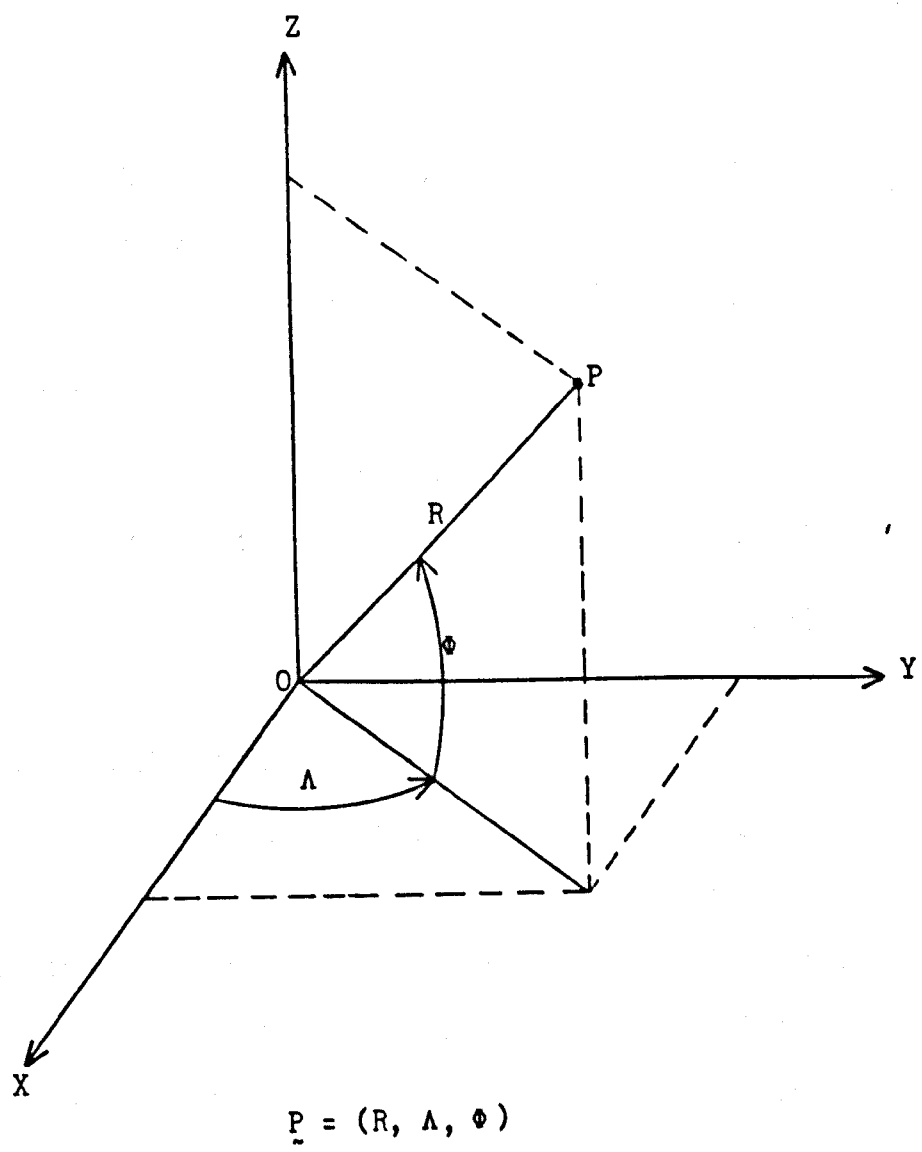


Fig B.II Spherical Polar Coordinates

ϕ : angle between OA and OP, measured from OA
towards the positive Z-direction (latitude).

The relationship between cartesian and spherical coordinates is given by

$$X = R \cos \phi \cos \Lambda \quad (\text{B.3})$$

$$Y = R \cos \phi \sin \Lambda \quad (\text{B.4})$$

$$Z = R \sin \phi \quad (\text{B.5})$$

The inverse relationship is

$$R = (X^2 + Y^2 + Z^2)^{\frac{1}{2}} \quad (\text{B.6})$$

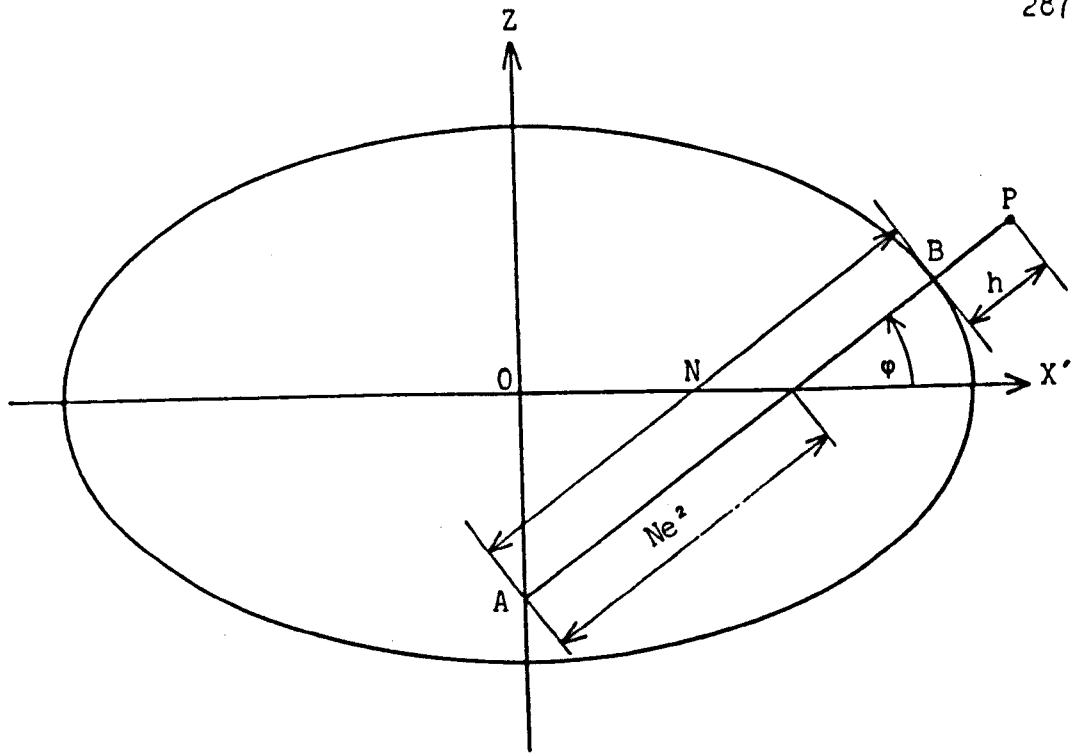
$$\Lambda = \tan^{-1} \frac{Y}{X} \quad (\text{B.7})$$

$$\phi = \sin^{-1} \frac{Z}{R} \quad (\text{B.8})$$

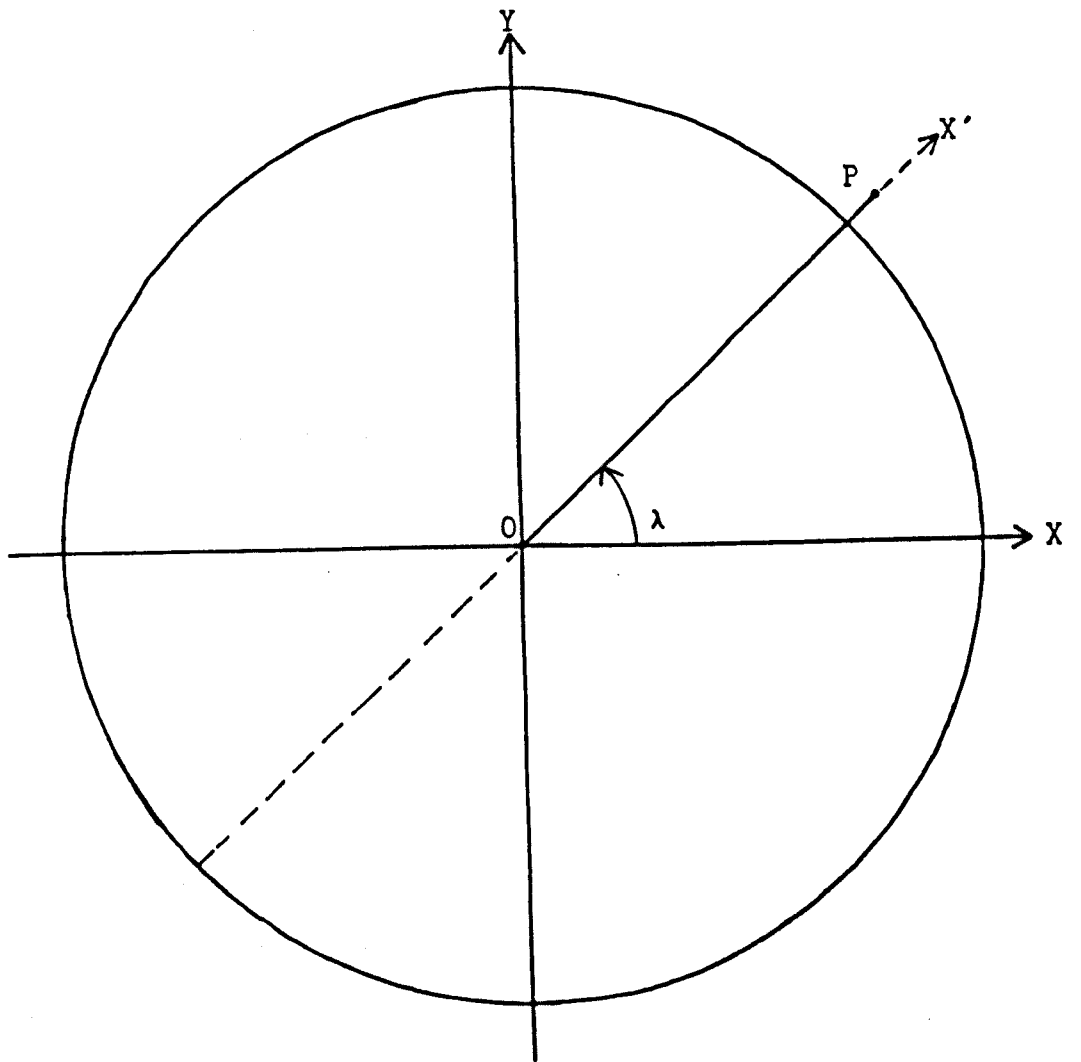
B.3 Spheroidal Representation

The reference surface that is most commonly adopted for geodetic computations is an oblate spheroid (an ellipsoid of rotation about its minor axis). This is such that the minor axis of the ellipsoid coincides with the Z - axis of the cartesian representation. A meridional plane (a plane which contains the minor axis and the point, P) is shown in fig. B.III. The line PA is at right angles to the tangent of the ellipsoid at B, the projection of P on the spheroid, and is called the normal at P.

The X - Y plane which includes the major axis of the ellipsoid is called the equatorial plane. The angle between the normal at P and the equatorial plane is the geodetic latitude, ϕ , as shown in fig.B.III. The angle between the meridional plane which includes the X - axis (eg the Greenwich Meridian) and the meridional plane



Meridional Section



Plan

Fig B.III Spheroidal Coordinates

at P, measured anti-clockwise from the X - axis, is the geodetic longitude of P, λ , and is illustrated in fig.B.III. The height, h, of P above the reference spheroid is the distance BP along the normal. The spheroidal representation of the position vector of P is

$$\underline{P} = (\varphi, \lambda, h) \quad (\text{B.9})$$

The equation of the spheroid in cartesian coordinates is

$$\frac{X^2}{a^2} + \frac{Y^2}{a^2} + \frac{Z^2}{b^2} = 1 \quad (\text{B.10})$$

where a : semi-major axis

b : semi-minor axis

The eccentricity, e, is given by

$$e^2 = \frac{a^2 - b^2}{a^2} \quad (\text{B.11})$$

and the flattening, f, by

$$f = \frac{a - b}{a} \quad (\text{B.12})$$

This gives

$$e^2 = 1 - (1 - f)^2 \quad (\text{B.13})$$

The prime vertical at P is the plane perpendicular to the meridional plane, which includes the normal. The distance AB is the radius of curvature (N) in the prime vertical and is given as (Bomford, 1980) :

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{\frac{1}{2}}} \quad (\text{B.14})$$

Referring to fig B.III, the transformation from geodetic to cartesian coordinates is given by

$$X = (N+h) \cos \varphi \cos \lambda \quad (\text{B.15})$$

$$Y = (N+h) \cos \varphi \sin \lambda \quad (\text{B.16})$$

$$Z = (N+h - Ne^2) \sin \varphi \quad (\text{B.17})$$

The reverse transformation is

$$\lambda = \tan^{-1} \frac{Y}{X} \quad (\text{B.18})$$

$$\varphi = \tan^{-1} \left[\frac{Z + Ne^2 \sin \varphi}{(X^2 + Y^2)^{\frac{1}{2}}} \right] \quad (\text{B.19})$$

$$h = \frac{X}{\cos \varphi \cos \lambda} - N \quad (\text{B.20})$$

Equation (B.19) can be evaluated iteratively by assuming an initial value of φ , which could be the geocentric latitude as given by (B.9).

It is frequently required to map small shifts in geodetic coordinates to corresponding cartesian displacements. These are obtained by differentiating (B.15), (B.16) and (B.17) and are given in matrix form as

$$\begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} = \begin{bmatrix} -\cos \lambda \cos \varphi & -\sin \lambda & \cos \lambda \cos \varphi \\ -\sin \lambda \sin \varphi & \cos \lambda & \sin \lambda \cos \varphi \\ \cos \varphi & 0 & \sin \varphi \end{bmatrix} \cdot \begin{bmatrix} (N+h) d\varphi \\ (N+h) \cos \varphi d\lambda \\ dh \end{bmatrix} \quad (\text{B.21})$$

Matrix S is orthogonal, so the reverse transformation is

$$D\mathbf{G} = \mathbf{S}^T D\mathbf{X} \quad (\text{B.22})$$

The values of $(N+h)d\varphi$, $(N+h)\cos\varphi d\lambda$ and dh , which are the components of \underline{DG} in (B.22), represent the changes in latitude, longitude and height, in units of length, due to corresponding changes dX , dY and dZ in the cartesian coordinates of P .

APPENDIX C

ORBIT DETERMINATION PARTIAL DERIVATIVES

APPENDIX C

ORBIT DETERMINATION PARTIAL DERIVATIVES

C.1 Evaluation of Satellite Acceleration due to Earth Attraction

The general component of the EF satellite acceleration vector due to the earth's attraction is given by (2.42). This can be written in matrix form as

$$\ddot{\mathbf{R}} = \mathbf{DS} \cdot \mathbf{DU} \quad (\text{C.1})$$

where $\ddot{\mathbf{R}} : (\ddot{X}, \ddot{Y}, \ddot{Z})^T$, the EF acceleration vector

$\mathbf{DU} : \left(\frac{\partial U}{\partial R}, \frac{\partial U}{\partial \Lambda}, \frac{\partial U}{\partial \Phi} \right)^T$, the first derivatives of the geopotential (equation (2.30)) with respect to the spherical coordinates of the satellite.

$\mathbf{DS} : \text{the } (3 \times 3) \text{ matrix of first derivatives of spherical w.r.t. cartesian coordinates, as given by}$

$$\mathbf{DS} = \begin{bmatrix} \frac{\partial R}{\partial X} & \frac{\partial \Lambda}{\partial X} & \frac{\partial \Phi}{\partial X} \\ \frac{\partial R}{\partial Y} & \frac{\partial \Lambda}{\partial Y} & \frac{\partial \Phi}{\partial Y} \\ \frac{\partial R}{\partial Z} & \frac{\partial \Lambda}{\partial Z} & \frac{\partial \Phi}{\partial Z} \end{bmatrix} \quad (\text{C.2})$$

The elements of \mathbf{DU} are derived by differentiating the expression for the geopotential in equation (2.30). They are :

$$\frac{\partial U}{\partial R} = \frac{-GM}{R^2} \left[1 + \sum_{n=2}^{\infty} \sum_{m=0}^n \left(\frac{a}{R} \right)^n (n+1) P_n^m (C_n^m \cos m\lambda + S_n^m \sin m\lambda) \right] \dots (C.3)$$

$$\frac{\partial U}{\partial \lambda} = \frac{GM}{R} \sum_{n=2}^{\infty} \sum_{m=0}^n \left(\frac{a}{R} \right)^n P_n^m (S_n^m \cos m\lambda - C_n^m \sin m\lambda) \dots (C.4)$$

$$\frac{\partial U}{\partial \phi} = \frac{GM}{R} \sum_{n=2}^{\infty} \sum_{m=0}^n \left(\frac{a}{R} \right)^n (P_n^{m+1} - m \tan \phi P_n^m) (C_n^m \cos m\lambda + S_n^m \sin m\lambda) \dots (C.5)$$

The elements of DS are computed by differentiating equations (B.6), (B.7) and (B.8). They are :

$$\frac{\partial R}{\partial X} = \frac{X}{R}, \quad \frac{\partial R}{\partial Y} = \frac{Y}{R}, \quad \frac{\partial R}{\partial Z} = \frac{Z}{R} \dots (C.6)$$

$$\frac{\partial \lambda}{\partial X} = \frac{-Y}{s^2}, \quad \frac{\partial \lambda}{\partial Y} = \frac{X}{s^2}, \quad \frac{\partial \lambda}{\partial Z} = 0 \dots (C.7)$$

$$\frac{\partial \phi}{\partial X} = \frac{-ZX}{sR^2}, \quad \frac{\partial \phi}{\partial Y} = \frac{-ZY}{sR^2}, \quad \frac{\partial \phi}{\partial Z} = \frac{1}{s} \left(1 - \frac{Z^2}{R^2} \right) \dots (C.8)$$

with

$$s = (X^2 + Y^2)^{\frac{1}{2}} \dots (C.9)$$

C.2 Partial Derivatives of Satellite Acceleration

The computation of the observation equation coefficients, discussed in Section 2.5.3, requires the partial derivatives of the IF satellite acceleration with respect to the state vector and force model unknowns. These derivatives are given in matrix form by equation (2.188) which is repeated here, retaining the same notation, as :

$$A_m = D1 \cdot X_m + D2 \cdot V_m + A_f \dots (C.10)$$

where

$$A_m = \begin{bmatrix} \frac{\partial \ddot{x}}{\partial x_o} & \dots & \frac{\partial \ddot{x}}{\partial \dot{z}_o} & \frac{\partial \ddot{x}}{\partial p_1} & \dots & \frac{\partial \ddot{x}}{\partial p_{np}} \\ \frac{\partial \ddot{y}}{\partial x_o} & & \cdot & \cdot & & \frac{\partial \ddot{y}}{\partial p_{np}} \\ \frac{\partial \ddot{z}}{\partial x_o} & & \cdot & \cdot & & \frac{\partial \ddot{z}}{\partial p_{np}} \end{bmatrix} \quad (C.11)$$

$$V_m = \begin{bmatrix} \frac{\partial \dot{x}}{\partial x_o} & \dots & \frac{\partial \dot{x}}{\partial \dot{z}_o} & \frac{\partial \dot{x}}{\partial p_1} & \dots & \frac{\partial \dot{x}}{\partial p_{np}} \\ \frac{\partial \dot{y}}{\partial x_o} & & \cdot & \cdot & & \frac{\partial \dot{y}}{\partial p_{np}} \\ \frac{\partial \dot{z}}{\partial x_o} & & \cdot & \cdot & & \frac{\partial \dot{z}}{\partial p_{np}} \end{bmatrix} \quad (C.12)$$

$$X_m = \begin{bmatrix} \frac{\partial x}{\partial x_o} & \dots & \frac{\partial x}{\partial \dot{z}_o} & \frac{\partial x}{\partial p_1} & \dots & \frac{\partial x}{\partial p_{np}} \\ \frac{\partial y}{\partial x_o} & & \cdot & \cdot & & \frac{\partial y}{\partial p_{np}} \\ \frac{\partial z}{\partial x_o} & & \cdot & \cdot & & \frac{\partial z}{\partial p_{np}} \end{bmatrix} \quad (C.13)$$

$$A_f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial \ddot{x}^*}{\partial p_1} & \dots & \frac{\partial \ddot{x}^*}{\partial p_{np}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial \ddot{y}^*}{\partial p_1} & \dots & \frac{\partial \ddot{y}^*}{\partial p_{np}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial \ddot{z}^*}{\partial p_1} & \dots & \frac{\partial \ddot{z}^*}{\partial p_{np}} \end{bmatrix} \quad (C.14)$$

$$D1 = \begin{bmatrix} \frac{\partial \ddot{x}}{\partial x} & \frac{\partial \ddot{x}}{\partial y} & \frac{\partial \ddot{x}}{\partial z} \\ \frac{\partial \ddot{y}}{\partial x} & \frac{\partial \ddot{y}}{\partial y} & \frac{\partial \ddot{y}}{\partial z} \\ \frac{\partial \ddot{z}}{\partial x} & \frac{\partial \ddot{z}}{\partial y} & \frac{\partial \ddot{z}}{\partial z} \end{bmatrix} \quad (C.15)$$

$$D2 = \begin{bmatrix} \frac{\partial \ddot{x}}{\partial \dot{x}} & \frac{\partial \ddot{x}}{\partial \dot{y}} & \frac{\partial \ddot{x}}{\partial \dot{z}} \\ \frac{\partial \ddot{y}}{\partial \dot{x}} & \frac{\partial \ddot{y}}{\partial \dot{y}} & \frac{\partial \ddot{y}}{\partial \dot{z}} \\ \frac{\partial \ddot{z}}{\partial \dot{x}} & \frac{\partial \ddot{z}}{\partial \dot{y}} & \frac{\partial \ddot{z}}{\partial \dot{z}} \end{bmatrix} \quad (C.16)$$

The acceleration partials, A_m , are integrated numerically to give the velocity and position partials, V_m and X_m , as functions of time, with the initial conditions of equations (2.189) and (2.190) (V_m^0 and X_m^0). To evaluate A_m , the matrices $D1$, $D2$ and A_f are required, and the rest of this appendix explains the method to compute them.

$D1$ is defined in (C.15) and it contains the partial derivatives of the satellite IF acceleration with respect to the components of the IF position vector. All the constituents of the force model, with the exception of along-track acceleration, are functions of position and therefore contribute to $D1$. By far the most important contribution, for near-earth satellites, is the earth's gravitational attraction (see Section 2.3.2). Equation (C.1) gives the satellite EF acceleration vector arising from the geopotential, U , as a

function of its EF coordinates. The required partial derivatives are then, in the EF reference frame,

$$D1E = \begin{bmatrix} \frac{\partial \ddot{X}}{\partial X} & \frac{\partial \ddot{X}}{\partial Y} & \frac{\partial \ddot{X}}{\partial Z} \\ \frac{\partial \ddot{Y}}{\partial X} & \frac{\partial \ddot{Y}}{\partial Y} & \frac{\partial \ddot{Y}}{\partial Z} \\ \frac{\partial \ddot{Z}}{\partial X} & \frac{\partial \ddot{Z}}{\partial Y} & \frac{\partial \ddot{Z}}{\partial Z} \end{bmatrix} \quad (C.17)$$

and these must be rotated to the IF, by

$$D1 = (PENQ)^T \cdot D1E \cdot (PENQ) \quad (C.18)$$

where P, E, N and Q are defined in Section 2.2.4.

Since the geopotential is given in terms of spherical coordinates (see equation (2.42)), the elements of D1E are computed as in the following example:

$$\begin{aligned} \frac{\partial \ddot{X}}{\partial X} &= \frac{\partial^2 U}{\partial X^2} = \frac{\partial}{\partial X} \left(\frac{\partial U}{\partial R} \frac{\partial R}{\partial X} + \frac{\partial U}{\partial \Lambda} \frac{\partial \Lambda}{\partial X} + \frac{\partial U}{\partial \Phi} \frac{\partial \Phi}{\partial X} \right) \\ &= \frac{\partial}{\partial X} \left(\frac{\partial U}{\partial R} \right) \cdot \frac{\partial R}{\partial X} + \frac{\partial}{\partial X} \left(\frac{\partial U}{\partial \Lambda} \right) \cdot \frac{\partial \Lambda}{\partial X} + \frac{\partial}{\partial X} \left(\frac{\partial U}{\partial \Phi} \right) \cdot \frac{\partial \Phi}{\partial X} \\ &\quad + \frac{\partial U}{\partial R} \frac{\partial^2 R}{\partial X^2} + \frac{\partial U}{\partial \Lambda} \frac{\partial^2 \Lambda}{\partial X^2} + \frac{\partial U}{\partial \Phi} \frac{\partial^2 \Phi}{\partial X^2} \dots \quad (C.19) \end{aligned}$$

where

$$\frac{\partial}{\partial X} \left(\frac{\partial U}{\partial R} \right) = \frac{\partial^2 U}{\partial R^2} \frac{\partial R}{\partial X} + \frac{\partial^2 U}{\partial R \partial \Lambda} \frac{\partial \Lambda}{\partial X} + \frac{\partial^2 U}{\partial R \partial \Phi} \frac{\partial \Phi}{\partial X} \dots \quad (C.20)$$

$$\frac{\partial}{\partial X} \left(\frac{\partial U}{\partial \Lambda} \right) = \frac{\partial^2 U}{\partial \Lambda \partial R} \frac{\partial R}{\partial X} + \frac{\partial^2 U}{\partial \Lambda^2} \frac{\partial \Lambda}{\partial X} + \frac{\partial^2 U}{\partial \Lambda \partial \Phi} \frac{\partial \Phi}{\partial X} \quad \dots \quad (C.21)$$

$$\frac{\partial}{\partial X} \left(\frac{\partial U}{\partial \Phi} \right) = \frac{\partial^2 U}{\partial \Phi \partial R} \frac{\partial R}{\partial X} + \frac{\partial^2 U}{\partial \Phi \partial \Lambda} \frac{\partial \Lambda}{\partial X} + \frac{\partial^2 U}{\partial \Phi^2} \frac{\partial \Phi}{\partial X} \quad \dots \quad (C.22)$$

There are similar expressions for all other elements of D1E, which can be expressed in matrix form as

$$D1E = DS \cdot D2U \cdot DS^T + U1 \cdot D2S1 + U2 \cdot D2S2 + U3 \cdot D2S3 \quad \dots \quad (C.23)$$

where DS: as defined in (C.2)

$$D2U = \begin{bmatrix} \frac{\partial^2 U}{\partial R^2} & \frac{\partial^2 U}{\partial R \partial \Lambda} & \frac{\partial^2 U}{\partial R \partial \Phi} \\ \frac{\partial^2 U}{\partial \Lambda \partial R} & \frac{\partial^2 U}{\partial \Lambda^2} & \frac{\partial^2 U}{\partial \Lambda \partial \Phi} \\ \frac{\partial^2 U}{\partial \Phi \partial R} & \frac{\partial^2 U}{\partial \Phi \partial \Lambda} & \frac{\partial^2 U}{\partial \Phi^2} \end{bmatrix} \quad \dots \quad (C.24)$$

$$U1 = \begin{bmatrix} \frac{\partial U}{\partial R} & \frac{\partial U}{\partial \Lambda} & \frac{\partial U}{\partial \Phi} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \dots \quad (C.25)$$

$$U2 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{\partial U}{\partial R} & \frac{\partial U}{\partial \Lambda} & \frac{\partial U}{\partial \Phi} \\ 0 & 0 & 0 \end{bmatrix} \quad \dots \quad (C.26)$$

$$U3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\partial U}{\partial R} & \frac{\partial U}{\partial \Lambda} & \frac{\partial U}{\partial \Phi} \end{bmatrix} \quad \dots \quad (C.27)$$

$$D2S1 = \begin{bmatrix} \frac{\partial^2 R}{\partial X^2} & \frac{\partial^2 R}{\partial X \partial Y} & \frac{\partial^2 R}{\partial X \partial Z} \\ \frac{\partial^2 \Lambda}{\partial X^2} & \frac{\partial^2 \Lambda}{\partial X \partial Y} & \frac{\partial^2 \Lambda}{\partial X \partial Z} \\ \frac{\partial^2 \Phi}{\partial X^2} & \frac{\partial^2 \Phi}{\partial X \partial Y} & \frac{\partial^2 \Phi}{\partial X \partial Z} \end{bmatrix} \quad \dots \quad (C.28)$$

$$D2S2 = \begin{bmatrix} \frac{\partial^2 R}{\partial Y \partial X} & \frac{\partial^2 R}{\partial Y^2} & \frac{\partial^2 R}{\partial Y \partial Z} \\ \frac{\partial^2 \Lambda}{\partial Y \partial X} & \frac{\partial^2 \Lambda}{\partial Y^2} & \frac{\partial^2 \Lambda}{\partial Y \partial Z} \\ \frac{\partial^2 \Phi}{\partial Y \partial X} & \frac{\partial^2 \Phi}{\partial Y^2} & \frac{\partial^2 \Phi}{\partial Y \partial Z} \end{bmatrix} \quad \dots \quad (C.29)$$

$$D2S3 = \begin{bmatrix} \frac{\partial^2 R}{\partial Z \partial X} & \frac{\partial^2 R}{\partial Z \partial Y} & \frac{\partial^2 R}{\partial Z^2} \\ \frac{\partial^2 \Lambda}{\partial Z \partial X} & \frac{\partial^2 \Lambda}{\partial Z \partial Y} & \frac{\partial^2 \Lambda}{\partial Z^2} \\ \frac{\partial^2 \Phi}{\partial Z \partial X} & \frac{\partial^2 \Phi}{\partial Z \partial Y} & \frac{\partial^2 \Phi}{\partial Z^2} \end{bmatrix} \quad \dots \quad (C.30)$$

Matrices $U1$, $U2$ and $U3$ are computed from equations (C.3), (C.4) and (C.5). $D2U$ is derived by differentiating each of these equations with respect to R , Λ and Φ .

For example :

$$\frac{\partial^2 U}{\partial \Phi^2} = \frac{GM}{R} \sum_{n=2}^{\infty} \sum_{m=0}^n \left(\frac{a}{R} \right)^n (C_n^m \cos m\Lambda + S_n^m \sin m\Lambda) \frac{\partial^2}{\partial \Phi^2} (P_n^m) \quad \dots \quad (C.31)$$

where

$$\begin{aligned} \frac{\partial^2}{\partial \Phi^2} (P_n^m) &= P_n^{m+2} - (m+1) \tan \Phi P_n^{m+1} - m \tan \Phi (P_n^{m+1} - m \tan \Phi P_n^m) \\ &\quad - m \sec^2 \Phi P_n^m \quad \dots \quad (C.32) \end{aligned}$$

Matrices D2S1, D2S2 and D2S3 are evaluated by differentiating (C.6), (C.7) and (C.8) with respect to the satellite EF coordinates, X, Y and Z. For example :

$$\frac{\partial^2 R}{\partial X^2} = \frac{\partial}{\partial X} \left(\frac{X}{R} \right) = \frac{1}{R} - \frac{X^2}{R^3} \quad \dots \quad (C.33)$$

$$\frac{\partial^2 R}{\partial X \partial Y} = \frac{\partial}{\partial Y} \left(\frac{X}{R} \right) = - \frac{XY}{R^3} \quad \dots \quad (C.34)$$

$$\frac{\partial^2 \Phi}{\partial X^2} = \frac{\partial}{\partial X} \left(- \frac{ZX}{SR^2} \right) = \frac{Z}{SR^2} \left(\frac{X^2}{S^2} + \frac{2X^2}{R^2} - 1 \right) \quad \dots \quad (C.35)$$

The contributions to D1 arising from the other constituents of the force model are easier to evaluate, since all other components are given in terms of the IF cartesian satellite position vector. However, for near-earth satellites it may prove unnecessary to include the contributions of other forces if it is found that their effect in the evaluation of the observation equation coefficients is insignificant (see Section 2.5.3).

The elements of matrix $D2$ are defined in (C.16) and they are obtained by differentiating the satellite IF acceleration vector with respect to each of the components of the velocity vector, $\dot{\mathbf{r}}$. $D2$ is non-zero only if drag, along-track and cross-track accelerations are present, since these are the only force model constituents which are functions of velocity. Even when these are present, $D2$ may be set to zero if this does not adversely change the rate of convergence of the least squares solution.

When the orbit determination unknowns include a number, np , of force model parameters, the matrix A_f has to be computed. As can be seen from (C.14), its first six columns, which correspond to the state vector unknowns, are all zero. Subsequent columns consist of the vectors

$$\frac{\partial \ddot{\mathbf{r}}^*}{\partial p_i} = \left(\frac{\partial \ddot{x}^*}{\partial p_i}, \frac{\partial \ddot{y}^*}{\partial p_i}, \frac{\partial \ddot{z}^*}{\partial p_i} \right)^T \quad (C.36)$$

and these are derived by directly differentiating the force model components which contain the parameter p_i .

For the determination of unknowns in the geopotential expansion of (2.30), the IF acceleration derivatives $\frac{\partial \ddot{\mathbf{r}}^*}{\partial C_n^m}$, $\frac{\partial \ddot{\mathbf{r}}^*}{\partial S_n^m}$, $\frac{\partial \ddot{\mathbf{r}}^*}{\partial GM}$ and

$\frac{\partial \ddot{\mathbf{r}}^*}{\partial a}$ may be required. Taking the derivatives with respect to the spherical harmonic coefficients C_n^m as an example, these are given in the EF reference frame by differentiating (C.1), as

$$\frac{\partial \ddot{\mathbf{r}}^*}{\partial C_n^m} = DS \cdot \frac{\partial}{\partial C_n^m} (DU) \quad (C.37)$$

where

$$\frac{\partial}{\partial C_n^m} (\nabla U) = \left(\frac{\partial^2 U}{\partial R \partial C_n^m}, \frac{\partial^2 U}{\partial \Lambda \partial C_n^m}, \frac{\partial^2 U}{\partial \Phi \partial C_n^m} \right)^T \quad (C.38)$$

The derivatives in (C.38) are given by differentiating (C.3), (C.4) and (C.5) with respect to C_n^m , as in

$$\frac{\partial^2 U}{\partial R \partial C_n^m} = -(n+1) \frac{GM}{R^2} \left(\frac{a}{R} \right)^n \cos m\lambda P_n^m \quad (C.39)$$

If the solution is in terms of normalized spherical harmonic coefficients, (C.39) becomes

$$\frac{\partial^2 U}{\partial R \partial \bar{C}_n^m} = N_n^m \frac{\partial^2 U}{\partial R \partial C_n^m} \quad (C.40)$$

where N_n^m is the normalizing factor defined in (2.39).

The EF acceleration partials must be rotated to the IF by

$$\frac{\partial \ddot{r}^*}{\partial \bar{C}_n^m} = (P \ E \ N \ Q)^T \frac{\partial \ddot{R}^*}{\partial \bar{C}_n^m} \quad (C.41)$$

The $\frac{\partial \ddot{r}^*}{\partial p_i}$ for other force model unknowns are computed more

easily. For example, for the solar radiation reflectance coefficient (equation 2.102),

$$\frac{\partial \ddot{r}^*}{\partial C_R} = \frac{I_0}{c} \left[\frac{A}{|r - r_j|} \right]^2 \left(\frac{A}{m} \right) \frac{r - r_j}{|r - r_j|} \quad (C.42)$$

and for the along-track acceleration (equation (2.91)),

$$\frac{\partial \ddot{r}^*}{\partial C_a} = \frac{\dot{r}}{r} \quad (C.43)$$

APPENDIX D

METHODS OF NUMERICAL INTEGRATION FOR ORBITS

APPENDIX D

METHODS OF NUMERICAL INTEGRATION FOR ORBITS

D.1 The Fourth Order Runge-Kutta Method

In this appendix the components of the IF acceleration, velocity and position vectors at time

$$t_i = t_0 + ih \quad (D.1)$$

where t_0 : starting epoch
 h : integration step size
 i : any integer

are denoted by

$$\ddot{\underline{r}}_i = (\ddot{x}_i, \ddot{y}_i, \ddot{z}_i) \quad (D.2)$$

$$\dot{\underline{r}}_i = (\dot{x}_i, \dot{y}_i, \dot{z}_i) = (u_i, v_i, w_i) \quad (D.3)$$

$$\underline{r}_i = (x_i, y_i, z_i) \quad (D.4)$$

The acceleration components are computed as functions of time, position and velocity, represented by (see equation (2.106))

$$\begin{aligned} \ddot{x}_i &= f_1(t_i, u_i, v_i, w_i, x_i, y_i, z_i) \\ \ddot{y}_i &= f_2(t_i, u_i, v_i, w_i, x_i, y_i, z_i) \\ \ddot{z}_i &= f_3(t_i, u_i, v_i, w_i, x_i, y_i, z_i) \end{aligned} \quad (D.5)$$

The velocity vector components are

$$u_i = \left(\frac{dx}{dt} \right)_{t=t_i} = f_4, \quad v_i = \left(\frac{dy}{dt} \right)_{t=t_i} = f_5, \quad w_i = \left(\frac{dz}{dt} \right)_{t=t_i} = f_6, \quad \dots \quad (D.6)$$

The full 4th order Runge-Kutta formulae for the double integration of the 2nd order differential equations (D.5) are :

$$\begin{aligned}
 u_{i+1} &= u_i + \frac{1}{6} (K_{11} + 2K_{12} + 2K_{13} + K_{14}) \\
 v_{i+1} &= v_i + \frac{1}{6} (K_{21} + 2K_{22} + 2K_{23} + K_{24}) \\
 w_{i+1} &= w_i + \frac{1}{6} (K_{31} + 2K_{32} + 2K_{33} + K_{34}) \\
 x_{i+1} &= x_i + \frac{1}{6} (K_{41} + 2K_{42} + 2K_{43} + K_{44}) \\
 y_{i+1} &= y_i + \frac{1}{6} (K_{51} + 2K_{52} + 2K_{53} + K_{54}) \\
 z_{i+1} &= z_i + \frac{1}{6} (K_{61} + 2K_{62} + 2K_{63} + K_{64})
 \end{aligned} \tag{D.7}$$

where

$$\begin{aligned}
 K_{11} &= hf_1(t_i, u_i, v_i, x_i, y_i, z_i) = h \dot{x}_i \\
 K_{21} &= h\dot{y}_i \\
 K_{31} &= h\dot{z}_i \\
 K_{41} &= hf_4 = hu_i \\
 K_{51} &= hv_i \\
 K_{61} &= hw_i \\
 K_{12} &= hf_1(t_i + \frac{h}{2}, u_i + \frac{K_{11}}{2}, v_i + \frac{K_{21}}{2}, \dots, z_i + \frac{K_{61}}{2}) \\
 K_{22} &= hf_2(t_i + \frac{h}{2}, u_i + \frac{K_{11}}{2}, v_i + \frac{K_{21}}{2}, \dots, z_i + \frac{K_{61}}{2}) \\
 K_{32} &= hf_3(t_i + \frac{h}{2}, \dots, z_i + \frac{K_{61}}{2}) \\
 K_{42} &= h(u_i + \frac{K_{11}}{2}) \\
 K_{52} &= h(v_i + \frac{K_{21}}{2}) \\
 K_{62} &= h(w_i + \frac{K_{31}}{2})
 \end{aligned}$$

$$K_{13} = hf_1 \left(t_i + \frac{h}{2}, u_i + \frac{K_{12}}{2}, v_i + \frac{K_{22}}{2}, \dots, z_i + \frac{K_{62}}{2} \right) \quad (D.8)$$

$$K_{23} = hf_2 \left(t_i + \frac{h}{2}, u_i + \frac{K_{12}}{2}, v_i + \frac{K_{22}}{2}, \dots, z_i + \frac{K_{62}}{2} \right)$$

$$K_{33} = hf_3 \left(t_i + \frac{h}{2}, \dots, z_i + \frac{K_{62}}{2} \right)$$

$$K_{43} = h \left(u_i + \frac{K_{12}}{2} \right)$$

$$K_{53} = h \left(v_i + \frac{K_{22}}{2} \right)$$

$$K_{63} = h \left(w_i + \frac{K_{32}}{2} \right)$$

$$K_{14} = hf_1 (t_i + h, u_i + K_{13}, v_i + K_{23}, \dots, z_i + K_{63})$$

$$K_{24} = hf_2 (t_i + h, u_i + K_{13}, v_i + K_{23}, \dots, z_i + K_{63})$$

$$K_{34} = hf_3 (t_i + h, \dots, z_i + K_{63})$$

$$K_{44} = h (u_i + K_{13})$$

$$K_{54} = h (v_i + K_{23})$$

$$K_{64} = h (w_i + K_{33})$$

The Runge-Kutta procedure starts with the initial satellite position and velocity vectors, \underline{r}_0 and $\dot{\underline{r}}_0$, and computes the next ones, \underline{r}_1 and $\dot{\underline{r}}_1$. The latter are now used to compute \underline{r}_2 and $\dot{\underline{r}}_2$ and so on, until sufficient values have been calculated for the Adams-Bashforth procedure to take over.

D.2 The Adams-Bashforth Predictor-Corrector Formulae

For simplicity, only the first components of the satellite position and velocity vectors, x_i and u_i , will be considered. The Adams-Bashforth predictor formulae for the numerical integration of (D.5) are:

$$\begin{aligned} u_{i+1}^* &= u_i + h b_0 \ddot{x}_i + h \sum_{j=1}^n b_j \nabla^j \ddot{x}_i \\ x_{i+1}^* &= x_i + h b_0 u_i + h \sum_{j=1}^n b_j \nabla^j u_i \end{aligned} \quad (D.9)$$

where ∇^j : j^{th} backward difference (see Section 2.4.2)

b_j : Adams-Bashforth coefficient for j^{th} difference

n : order of the predictor-corrector method.

The coefficients, b_j , are derived from

$$b_n + \frac{1}{2} b_{n-1} + \frac{1}{3} b_{n-2} + \dots + \frac{1}{n+1} b_0 = 1 \quad (D.10)$$

giving $b_0 = 1$, $b_1 = \frac{1}{2}$, $b_2 = \frac{5}{12}$ etc.

The predicted acceleration vector is now evaluated from

$$\begin{aligned} \ddot{x}_{i+1}^* &= f_1(t_{i+1}, u_{i+1}^*, v_{i+1}^*, w_{i+1}^*, x_{i+1}^*, y_{i+1}^*, z_{i+1}^*) \\ &\dots \end{aligned} \quad (D.11)$$

and the corrector formulae are

$$\begin{aligned} u_{i+1}^{**} &= u_i + h c_0 \ddot{x}_{i+1}^* + h \sum_{j=1}^n c_j \nabla^j \ddot{x}_{i+1}^* \\ x_{i+1}^{**} &= x_i + h c_0 u_{i+1}^* + h \sum_{j=1}^n c_j \nabla^j u_{i+1}^* \end{aligned} \quad (D.12)$$

The coefficients , c_j , can be derived from

$$c_n + \frac{1}{2} c_{n-1} + \frac{1}{3} c_{n-2} + \dots + \frac{1}{n+1} c_0 = \begin{cases} 1 & n=0 \\ 0 & n=1,2,3,\dots \end{cases} \quad \dots \quad (D.13)$$

resulting in $c_0 = 1$, $c_1 = \frac{-1}{2}$, $c_2 = \frac{-1}{12}$ etc.

Equations (D.9) for the predictor formulae can be expressed in ordinate form for any order,n, as

$$u_{i+1}^* = u_i + h \sum_{j=0}^n \beta_j \ddot{x}_{i-j} \quad (D.14)$$

$$x_{i+1}^* = x_i + h \sum_{j=0}^n \beta_j u_{i-j}$$

and the coefficients β_j are given by (Henrici, 1962)

$$\beta_j = (-1)^j \sum_{\ell=j}^n \binom{\ell}{j} b_\ell \quad (D.15)$$

where b_ℓ : coefficients from (D.9)

$\binom{\ell}{j}$: number of combinations of ℓ things taken j at a time, computed from

$$\binom{\ell}{j} = \frac{\ell!}{j! (\ell-j)!} \quad (D.16)$$

Similarly, equations (D.12) for the corrector can be written as :

$$\begin{aligned}
 u_{i+1}^{**} &= u_i + h\gamma_0 \ddot{x}_{i+1}^* + h \sum_{j=1}^n \gamma_j \ddot{x}_{i-j+1} \\
 x_{i+1}^{**} &= x_i + h\gamma_0 u_{i+1}^* + h \sum_{j=1}^n \gamma_j u_{i-j+1}
 \end{aligned}
 \tag{D.17}$$

where

$$\gamma_j = (-1)^j \sum_{\ell=j}^n \binom{\ell}{j} c_\ell
 \tag{D.18}$$

c_ℓ : coefficients from (D.12)

These ordinate forms of (D.14) and (D.17) are easier to evaluate because there is no need to compute a difference table every time a new value is obtained.

APPENDIX E

POLYNOMIAL REPRESENTATION AND INTERPOLATION

APPENDIX E

POLYNOMIAL REPRESENTATION AND INTERPOLATION

E.1 Chebyshev Polynomials

A function of a variable, t , can be represented in terms of Chebyshev polynomials in the interval from t_i to $t_i + \Delta t$, as :

$$f(t) = \sum_{k=1}^n \alpha_k \cos k \theta \quad (E.1)$$

where

$$\theta = \cos^{-1} \left[\frac{(t - t_i) - \frac{\Delta t}{2}}{\Delta t} \right] \quad (E.2)$$

and n : order of the polynomial.

The coefficients α_k are computed by evaluating the function at $n+1$ specific data points, $f(t_j)$, which are given by

$$\theta_j = \left(\frac{2j+1}{n+1} \right) \frac{\pi}{2} \quad j = 0, 1, \dots, n \quad (E.3)$$

where

$$\theta_j = \cos^{-1} \left[\frac{(t_j - t_i) - \frac{\Delta t}{2}}{\Delta t} \right] \quad (E.4)$$

The Chebyshev coefficients are

$$\alpha_0 = \frac{1}{n+1} \sum_{j=0}^n f(t_j) \quad (E.5)$$

$$\alpha_k = \frac{2}{n+1} \sum_{j=0}^n f(t_j) \cos k \theta_j \quad (E.6)$$

In practice, (E.1) can be truncated after ℓ ($< n$) terms, such that

$$f(t) = \sum_{k=1}^{\ell} \alpha_k \cos k \theta \quad (E.7)$$

E.2 Everett Interpolation

A function, $f(t)$, which is given at discrete data points at

$$\dots, t_{i-2}, t_{i-1}, t_i, t_{i+1}, t_{i+2}, \dots$$

where

$$t_{i+m} = t_i + m\Delta t \quad (\text{E.8})$$

can be interpolated in the interval from t_i to t_{i+1} by using the Everett central difference formula

$$\begin{aligned} f(t) = & E_0 f(t_i) + E_2 \delta^2 f(t_i) + E_4 \delta^4 f(t_i) + \dots \\ & + F_0 f(t_{i+1}) + F_2 \delta^2 f(t_{i+1}) + F_4 \delta^4 f(t_{i+1}) + \dots \end{aligned} \quad \dots (\text{E.9})$$

with

$$u = \frac{t - t_i}{\Delta t} \quad (\text{E.10})$$

and

$$\begin{aligned} E_0 &= 1 - u \\ E_2 &= u(1-u)(2-u)/3! \\ E_4 &= (-1-u)u(1-u)(2-u)(3-u)/5! \end{aligned} \quad (\text{E.11})$$

$$\begin{aligned} F_0 &= u \\ F_2 &= u(u^2 - 1)/3! \\ F_4 &= u(u^2 - 1)(u^2 - 4)/5! \end{aligned} \quad (\text{E.12})$$

The term $\delta^n f(t_i)$ is the n^{th} central difference of $f(t_i)$, such that

$$\delta^2 f(t_i) = f(t_{i-1}) - 2f(t_i) + f(t_{i+1}) \quad (\text{E.13})$$

$$\begin{aligned} \delta^4 f(t_i) = & f(t_{i-2}) - 4f(t_{i-1}) + 6f(t_i) - 4f(t_{i+1}) \\ & + f(t_{i+2}) \end{aligned} \quad (\text{E.14})$$

When (E.9) is truncated after the n^{th} (even) central difference, then $n+2$ discrete data points must be available, equally balanced about the relevant interval.

APPENDIX F

WEIGHTED LEAST SQUARES

APPENDIX F

WEIGHTED LEAST SQUARES

F.1 Derivation of Normal Equations

Given the set of observation equations

$$Ax = b + v \quad (F.1)$$

which are weighted to give

$$W^{\frac{1}{2}}Ax = W^{\frac{1}{2}}b + W^{\frac{1}{2}}v \quad (F.2)$$

the object is to obtain the most probable value of vector x .

This is the value of x which minimises the sum of the squares of the weighted residuals, i.e.

$$v^T W v = \text{minimum} \quad (F.3)$$

This gives

$$\frac{\partial(v^T W v)}{\partial x} = 0 \quad (F.4)$$

From (F.2):

$$W^{\frac{1}{2}}v = W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}b \quad (F.5)$$

Squaring gives

$$v^T W v = (W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}b)^T \cdot (W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}b) \quad (F.6)$$

and multiplying out the bracketed terms, gives

$$v^T W v = x^T A^T W A x - x^T A^T W b - b^T W A x + b^T W b \quad (F.7)$$

Differentiating (F.7) with respect to x gives, for a minimum,

$$\frac{\partial(v^T W v)}{\partial x} = 2A^T W A x - A^T W b - A^T W b + 0 = 0 \quad (F.8)$$

So, for a minimum,

$$A^T W A x = A^T W b \quad (F.9)$$

and equation (F.9) is the matrix form of the normal equations.

F.2 Symmetric and Positive - Definite Properties of the Normal Matrix

F.2.1 Symmetry

The normal coefficient matrix is given by

$$N = A^T W A \quad (F.10)$$

The transpose is

$$N^T = (A^T W A)^T = A^T W^T A = A^T W A = N$$

since W is symmetric.

F.2.2 Positive-Definiteness

Any matrix, A , of order $(n \times n)$ is positive-definite if the quadratic form

$$y^T A y > 0$$

$$\text{where } y^T = [y_1, y_2, \dots, y_n]$$

and y is any real non-zero vector.

For the normal coefficient matrix, given by

$$N = A^T W A$$

the quadratic form is

$$\begin{aligned} y^T N y &= y^T \cdot (A^T W A) \cdot y \\ &= y^T \cdot (A^T W^{\frac{1}{2}} \cdot W^{\frac{1}{2}} A) \cdot y \\ &= y^T \cdot (W^{\frac{1}{2}} A)^T \cdot (W^{\frac{1}{2}} A) y \\ &= (W^{\frac{1}{2}} A y)^T \cdot (W^{\frac{1}{2}} A y) > 0 \end{aligned}$$

The above result is true since it is the sum of the squares of the vector $(W^{\frac{1}{2}}Ay)$.

The positive-definiteness of the normal equation matrix is the property which makes possible the solution of the normal equations by using Choleski's method of triangular decomposition, which is described in F.3.. This method involves obtaining the square roots of the diagonal elements of the normal coefficient matrix and would fail if any of these had values less than or equal to zero. The fact that the normal coefficient matrix, N , is positive-definite implies that the diagonal elements are positive, since

$$y^T N y > 0$$

for any real, non-zero vector y .

Choosing

$$y^T = [1, 0, 0, \dots, 0]$$

gives

$$d_1 > 0$$

where d_1 is the first diagonal element of matrix N . Similarly, choosing other suitable vectors y , gives

$$d_2, d_3, \dots, d_n > 0$$

Thus the elements on the leading diagonal of N are all positive.

F.3 Choleski's Method of Symmetric Decomposition

Choleski's method of symmetric decomposition can be used to solve the normal equations in a least squares network adjustment in the three steps outlined below :

1. Decomposition of the normal matrix N into

$$N = L L^T \quad (F.11)$$

where L and L^T are lower and upper triangular matrices respectively. An example of this decomposition for a (4×4) matrix is given below:

$$N = L \cdot L^T$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \cdot \begin{bmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{bmatrix}$$

By matrix multiplication this gives :

$$\begin{aligned} a_{11} &= l_{11}^2 & \therefore l_{11} &= \sqrt{a_{11}} \\ a_{12} &= l_{11} \cdot l_{21} & \therefore l_{21} &= a_{12}/l_{11} \\ a_{13} &= l_{11} \cdot l_{31} & \therefore l_{31} &= a_{13}/l_{11} \\ a_{14} &= l_{11} \cdot l_{41} & \therefore l_{41} &= a_{14}/l_{11} \\ a_{22} &= l_{21}^2 + l_{22}^2 & \therefore l_{22} &= \sqrt{(a_{22} - l_{21}^2)} \\ a_{23} &= l_{21} \cdot l_{31} + l_{22} \cdot l_{32} & \therefore l_{32} &= (a_{23} - l_{21} \cdot l_{31})/l_{22} \end{aligned}$$

$$a_{24} = l_{21} \cdot l_{41} + l_{22} \cdot l_{42} \quad \therefore \quad l_{42} = (a_{24} - l_{21} \cdot l_{41}) / l_{22}$$

$$a_{33} = l_{31}^2 + l_{32}^2 + l_{33}^2 \quad \therefore \quad l_{33} = \sqrt{(a_{33} - l_{31}^2 - l_{32}^2)}$$

$$a_{34} = l_{31} \cdot l_{41} + l_{32} \cdot l_{42} + l_{33} \cdot l_{43} \quad \therefore \quad l_{43} = (a_{34} - l_{31} \cdot l_{41} - l_{32} \cdot l_{42}) / l_{33}$$

$$a_{44} = l_{41}^2 + l_{42}^2 + l_{43}^2 + l_{44}^2 \quad \therefore \quad l_{44} = \sqrt{(a_{44} - l_{41}^2 - l_{42}^2 - l_{43}^2)}$$

The general elements are given by the algorithms

$$l_{jj} = \sqrt{(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2)} \quad \dots (F.12)$$

$$l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} \cdot l_{jk}) / l_{jj} \quad (i > j) \dots (F.13)$$

2. Forward Substitution

The normal equations can be written as

$$L(L^T x) = d \quad \dots (F.14)$$

$$\text{or,} \quad Lf = d \quad \dots (F.15)$$

The vector, f , is to be determined by forward substitution. Thus, for the (4×4) matrix:

$$L \cdot f = d$$

$$\begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$$

and

$$d_1 = l_{11} \cdot f_1 \quad \therefore f_1 = d_1 / l_{11}$$

$$d_2 = l_{21} \cdot f_1 + l_{22} \cdot f_2 \quad \therefore f_2 = (d_2 - l_{21} \cdot f_1) / l_{22}$$

$$d_3 = l_{31} \cdot f_1 + l_{32} \cdot f_2 + l_{33} \cdot f_3 \quad \therefore f_3 = (d_3 - l_{31} \cdot f_1 - l_{32} \cdot f_2) / l_{33}$$

$$d_4 = l_{41} \cdot f_1 + l_{42} \cdot f_2 + l_{43} \cdot f_3 + l_{44} \cdot f_4 \quad \therefore f_4 = (d_4 - l_{41} \cdot f_1 - l_{42} \cdot f_2 - l_{43} \cdot f_3) / l_{44}$$

The general element, f_i , can be expressed as

$$f_i = (d_i - \sum_{k=1}^{i-1} l_{ik} f_k) / l_{ii} \quad \dots (F.16)$$

3. Back Substitution

This determines the vector of unknowns, x , from the equation

$$L^T x = f \quad \dots (F.17)$$

where f was determined in step 2 above. Back substitution gives

$$L^T \cdot x = f$$

$$\begin{bmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$$

and

$$f_4 = l_{44} \cdot x_4 \quad \therefore x_4 = f_4 / l_{44}$$

$$f_3 = l_{33} \cdot x_3 + l_{43} \cdot x_4 \quad \therefore x_3 = (f_3 - l_{43} \cdot x_4) / l_{33}$$

$$f_2 = l_{22} \cdot x_2 + l_{32} \cdot x_3 + l_{42} \cdot x_4 \quad \therefore x_2 = (f_2 - l_{32} \cdot x_3 - l_{42} \cdot x_4) / l_{22}$$

$$f_1 = l_{11} \cdot x_1 + l_{21} \cdot x_2 + l_{31} \cdot x_3 + l_{41} \cdot x_4 \quad \therefore x_1 = (f_1 - l_{21} \cdot x_2 - l_{31} \cdot x_3 - l_{41} \cdot x_4) / l_{11}$$

The ground element, x_i , is given by

$$x_i = (f_i - \sum_{k=i+1}^n l_{ki} x_k) / l_{ii} \quad (F.18)$$

The algorithms given in formulae (F.11) to (F.18) can be utilized in a computer program for both solving the normal equations and finding the inverse of the normal coefficient matrix. An important storage saving characteristic is that once an element, l_{ij} , has been determined in the symmetric decomposition of N into L and L^T , the corresponding element, a_{ij} , is no longer required, so that the program need only operate on one matrix and replace old values by new ones.

APPENDIX G

COMPUTATION OF GPS TIME AND SATELLITE COORDINATES

APPENDIX G

COMPUTATION OF GPS TIME AND SATELLITE COORDINATES

G.1 Correction of Satellite Time to GPS Time

A pseudo-range measurement is the range equivalent of the difference between the local time of reception and the satellite time of transmission of a GPS signal. The satellite clock is not synchronized with GPS time and so the satellite time of transmission must be corrected to the GPS time of transmission by

$$t_t = t_s - \Delta t_s \quad (G.1)$$

where t_t : GPS time of transmission in seconds

t_s : satellite time of transmission in seconds

Δt_s : correction computed from parameters in the satellite data message (see Section 5.3.4).

The correction term, Δt_s , is given by

$$\Delta t_s = a_{f0} + a_{f1} (t_s - t_{oc}) + a_{f2} (t_s - t_{oc})^2 + \Delta t_{re} \quad (G.2)$$

where a_{f0} , a_{f1} , a_{f2} : polynomial coefficients given in subframe 1 of the data message

t_{oc} : clock reference time in seconds, also given in the data message

Δt_{re} : relativistic correction to the satellite clock (see Section 5.4.5.3)

To account for beginning or end of week crossovers, if the term $(t_s - t_{oc})$ in equation (G.2) exceeds 302400 seconds then 604800 (1 week) must be subtracted from it. Alternatively, if it is less than -302400 seconds then 604800 must be added to it.

The relativistic correction, Δt_{re} , is given by (ICD-GPS-200, 1981)

$$\Delta t_{re} = F e A^{\frac{1}{2}} \sin E_k \quad (G.3)$$

where

$$F = -4.443 \times 10^{-10} \text{ sm}^{-\frac{1}{2}} \quad (G.4)$$

The eccentricity, e , and semi-major axis, A , in equation (G.3) are given in the satellite data message. The computation of the eccentric anomaly, E_k , is explained in the next section.

A parameter called age of data clock (AODC) is transmitted in the navigation message in order to enable the user to assess the quality of the satellite clock data. This gives the time, t_{LC} , of the last tracking observation which was used to compute the satellite clock parameters, as

$$\text{AODC} = t_{oc} - t_{LC} \quad (G.5)$$

G.2 Satellite Cartesian Coordinates from Orbital Elements

The orbital elements which are transmitted in subframes 2 and 3 of the satellite navigation message (see Section 5.3.4) are listed in fig G.I. These are used in order to compute the WGS-72 satellite cartesian coordinates. The instantaneous position of the satellite, referred to the WGS-72 coordinate axes (X, Y, Z), is

SYMBOL	EXPLANATION
M_0	Mean anomaly at reference time (radians)
Δn	Mean motion difference from computed value (rad s^{-1})
e	Eccentricity
$A^{\frac{1}{2}}$	Square root of semi-major axis ($\text{m}^{\frac{1}{2}}$)
Ω_0	Right ascension of ascending node at reference time (radians)
i_0	Inclination at reference time (radians)
ω	Argument of perigee (radians)
IDOT	Rate of change of inclination (rad s^{-1})
$\dot{\Omega}$	Rate of right ascension (rad s^{-1})
C_{uc}	Amplitude of cosine correction term to the argument of latitude (radians)
C_{us}	Amplitude of the sine correction term to the argument of latitude (radians)
C_{rc}	Amplitude of the cosine correction term to the geocentric radius (metres)
C_{rs}	Amplitude of the sine correction term to the geocentric radius (metres)
C_{ic}	Amplitude of the cosine correction term to the inclination (radians)
C_{is}	Amplitude of the sine correction term to the inclination (radians)

Fig G.I Ephemeris Parameters in Satellite Data Message

described by 3 angles and a distance, as shown in fig G.II.

These are:

Ω_k : longitude of the ascending node of the instantaneous orbit (radians)

i_k : inclination of the instantaneous orbital plane (radians)

u_k : argument of latitude of satellite in orbital plane (radians)

r_k : geocentric radius to satellite (metres)

Before the satellite cartesian coordinates can be evaluated, the 4 parameters listed above have to be computed from the elements in fig G.I. The computation of Ω_k , i_k , u_k and r_k is explained in the following steps:

(i) satellite mean motion:

$$n_o = \left(\frac{GM}{A^3} \right)^{\frac{1}{2}} \quad (G.6)$$

where n_o : mean motion in rad s^{-1}

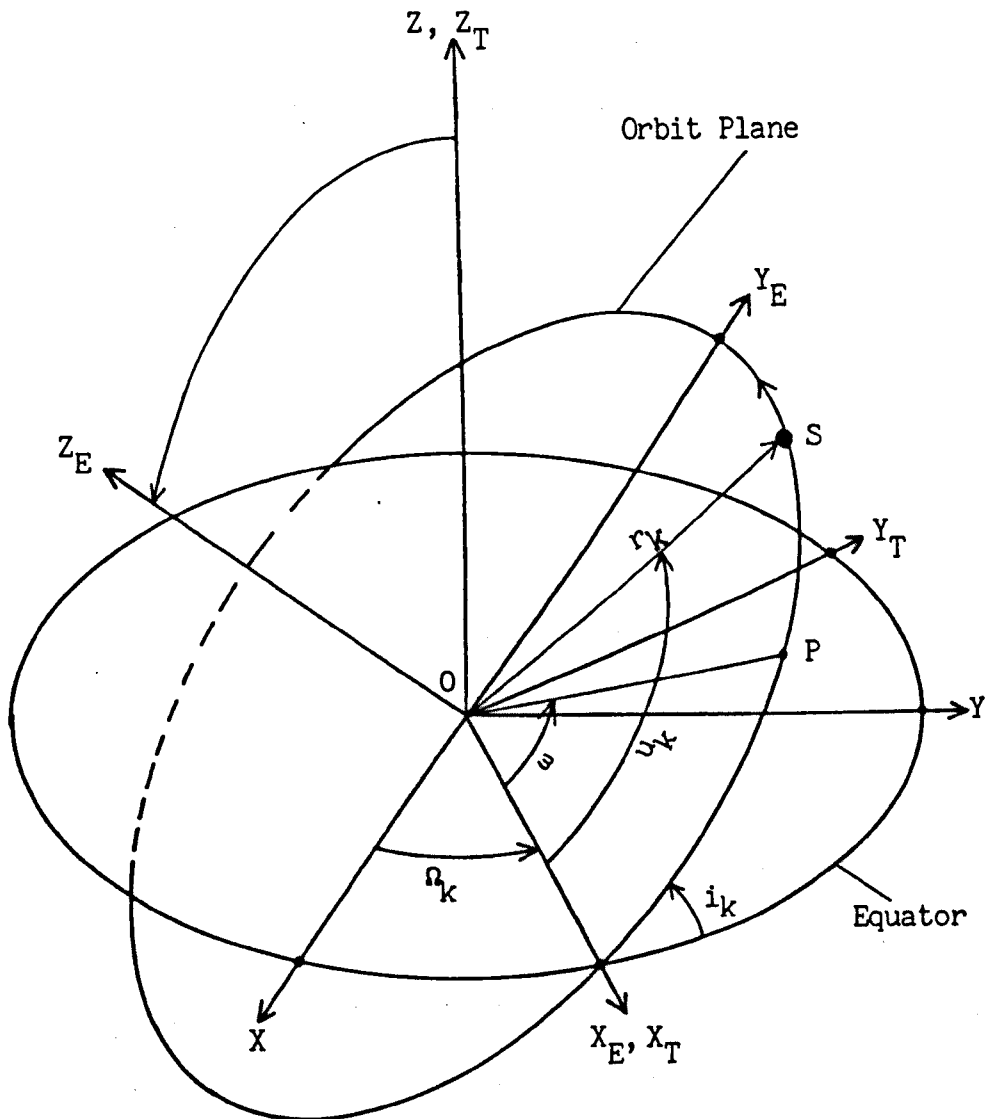
GM : WGS-72 value of geocentric gravitational constant ($3.986008 \times 10^{14} \text{ m}^3 \text{ s}^{-2}$)

(ii) GPS time from reference epoch:

$$t_k = t_t - t_{oe} \quad (G.7)$$

where t_t : GPS time of transmission in seconds from start of week (see Section G.1)

t_k : time elapsed from reference epoch, t_{oe} , in seconds



N : ascending node

P : perigee

S : satellite

O : geocentre

Fig G.II Instantaneous Satellite Orbit

To account for beginning or end of week crossovers, if t_k is greater than 302400 then 604800 must be subtracted from it, and if it is less than - 302400 then 604800 must be added to it.

(iii) correct mean motion :

$$n = n_o + \Delta n \quad (G.8)$$

where n : correct mean motion in rad s^{-1}

(iv) mean anomaly :

$$M_k = M_o + n t_k \quad (G.9)$$

where M_k : mean anomaly (radians) at time of transmission

(v) eccentric anomaly :

$$E_k = M_k + e \sin E_k \quad (G.10)$$

where E_k : eccentric anomaly in radians (see fig G.III)

Equation (G.10) is Kepler's equation and must be solved iteratively. The simplest method is to make an initial approximation of E_k , $E_k^{(0)}$, at the value of M_k .

A better approximation, $E_k^{(1)}$, is then

$$E_k^{(1)} = M_k + e \sin E_k^{(0)} \quad (G.11)$$

$E_k^{(1)}$ is similarly used in place of $E_k^{(0)}$ to derive $E_k^{(2)}$,

and so on until successive approximations agree to a

specified limit. The above method works well with the

GPS nominally circular (small eccentricity) orbits.

(vi) true anomaly :

$$\cos v_k = \frac{\cos E_k - e}{1 - e \cos E_k} \quad (G.12)$$

$$\sin v_k = \frac{(1-e^2)^{\frac{1}{2}} \sin E_k}{1 - e \cos E_k} \quad (G.13)$$

where v_k : true anomaly (radians)

The physical meaning of true anomaly is illustrated in fig G.III. The proofs of equations (G.12) and (G.13) are given in Section G.3.

(vii) provisional argument of latitude :

$$\varphi_k = v_k + \omega \quad (G.14)$$

where φ_k : provisional argument of latitude (radians)

(viii) corrected argument of latitude (see fig G.II):

$$u_k = \varphi_k + \delta u_k \quad (G.15)$$

where

$$\delta u_k = C_{us} \sin 2 \varphi_k + C_{uc} \cos 2 \varphi_k \quad (G.16)$$

(ix) provisional geocentric radius :

$$r = A(1 - e \cos E_k) \quad (G.17)$$

where r : provisional geocentric radius in metres

(see fig G.III)

A proof of equation (G.17) is given in the next section.

(x) correct geocentric radius (see fig G.II):

$$r_k = r + \delta r_k \quad (G.18)$$

where

$$\delta r_k = C_{rs} \sin 2 \varphi_k + C_{rc} \cos 2 \varphi_k \dots (G.19)$$

(xi) provisional inclination:

$$i = i_o + (I \text{ DOT}) t_k \quad (G.20)$$

where i : provisional value of the inclination of
the orbital plane (radians)

(xii) correct inclination (see fig G.II):

$$i_k = i + \delta i_k \quad (G.21)$$

where

$$\delta i_k = C_{is} \sin 2 \varphi_k + C_{ic} \cos 2 \varphi_k \quad (G.22)$$

(xiii) longitude of the ascending node (see fig G.II):

$$\Omega_k = \Omega_o + (\dot{\Omega} - \dot{\Omega}_e) t_k - \dot{\Omega}_e t_{oe} \quad (G.23)$$

where $\dot{\Omega}_e$: WGS-72 value of the earth's rotation
rate (7.292115147 rad s⁻¹)

From the instantaneous satellite parameters, Ω_k , i_k , u_k , and r_k , the next step is to compute the satellite cartesian coordinates in the plane of the orbital ellipse, referred to the axes X_E , Y_E , and Z_E (see fig G.II). These are given as:

$$\begin{aligned}
 X_E &= r_k \cos u_k \\
 Y_E &= r_k \sin u_k \\
 Z_E &= 0
 \end{aligned}
 \tag{G.24}$$

The coordinates of equations (G.24) are then rotated through a clockwise angle, i_k , about the X_E -axis in order to bring the Z_E - axis into coincidence with the WGS-72 Z -axis (see fig G.II). This rotation is expressed as :

$$\begin{aligned}
 X_T &= X_E \\
 Y_T &= Y_E \cos i_k \\
 Z_T &= Y_E \sin i_k
 \end{aligned}
 \tag{G.25}$$

The final transformation is a clockwise rotation through Ω_k radians about the Z_T -axis, which gives the WGS-72 satellite cartesian coordinates :

$$\begin{aligned}
 X &= X_T \cos \Omega_k - Y_T \sin \Omega_k \\
 Y &= Y_T \sin \Omega_k + X_T \cos \Omega_k \\
 Z &= Z_T
 \end{aligned}
 \tag{G.26}$$

This section has outlined the steps in the computation of the WGS-72 cartesian satellite coordinates from the transmitted orbital elements of fig G.I. The only variable in the above process is the GPS time of transmission, t_t . All other parameters are derived from the orbital elements, using t_t .

All angular quantities in this appendix have been given in radians. However, most of the ephemeris values are in semi-circle units (ICD-GPS-200, 1981). These are converted to radians by multiplication with π , where

$$\pi = 3.1415926535898 \quad (\text{G.27})$$

To enable the GPS user to assess the quality of the satellite orbital elements, an age of data ephemeris (AODE) parameter is transmitted. This is similar to AODC, described in the previous section, and gives the GPS time T_{Le} , of the last tracking measurement that was used in the determination of the satellite ephemeris, as:

$$\text{AODE} = t_{oe} - t_{Le} \quad (\text{G.28})$$

The Control Segment will ensure that the AODE value shall equal the AODC value defined in Section G.1.

G.3 Proofs of relationship Between Orbit Semi-major Axis and Geocentric Radius and Between Eccentric and True Anomalies

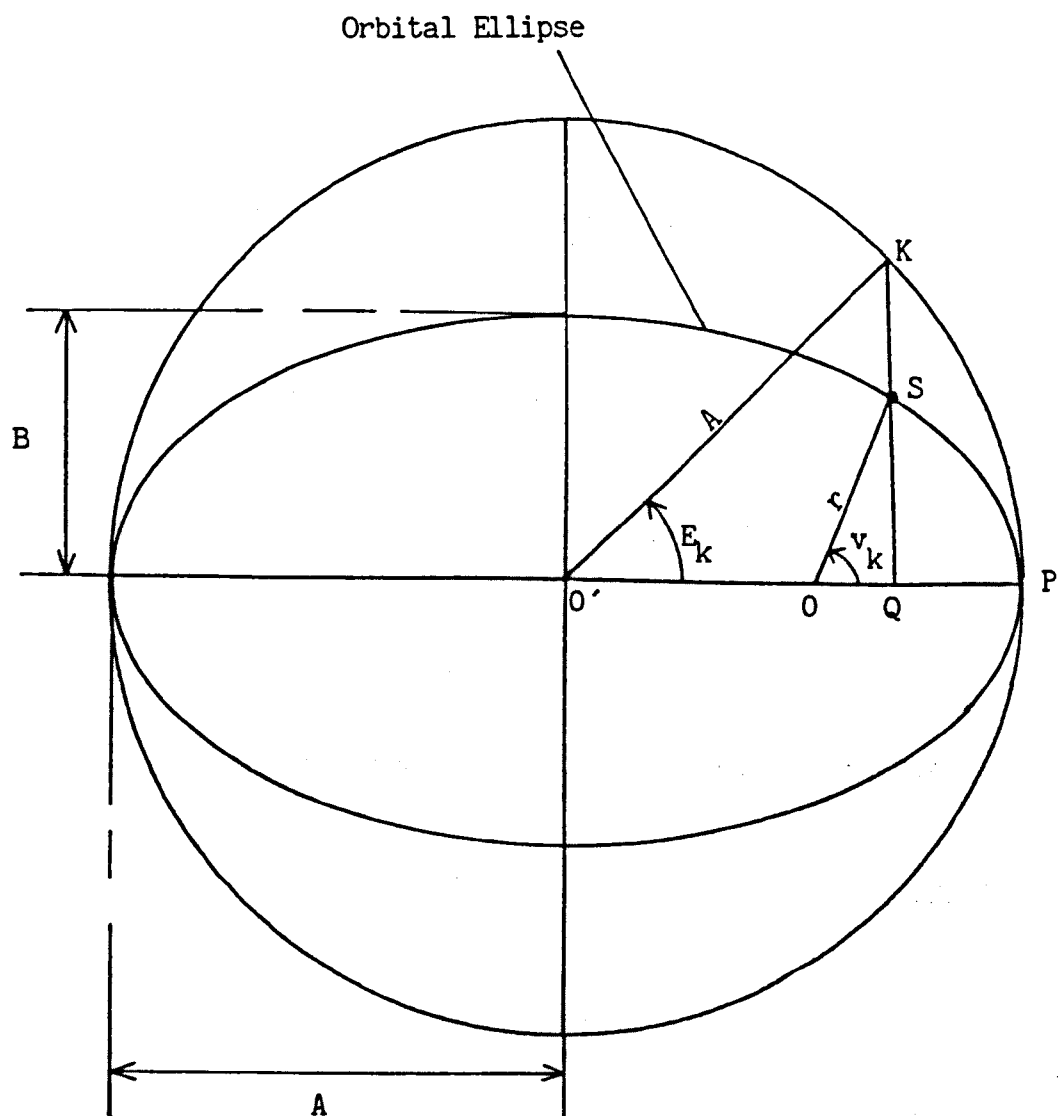
The mean orbital ellipse is illustrated in fig G.III.

From the geometry of the ellipse:

$$OQ = O'Q - Ae \quad (\text{G.29})$$

This gives:

$$r \cos v_k = A \cos E_k - Ae \quad (\text{G.30})$$



O : geocentre

P : perigee

S : satellite

A : semi-major axis

B : semi-minor axis

E_k : eccentric anomaly

v_k : true anomaly

Fig G.III Eccentric and True Anomalies

where the eccentric anomaly, E_k , true anomaly, v_k , and semi-major axis, A , are as labelled in fig G.III. Also from geometry:

$$QS = QK \times \frac{B}{A} \quad (G.31)$$

where B : ellipse semi-minor axis

Equation (G.31) can be easily verified. The equation of the ellipse gives :

$$\frac{O'Q^2}{A^2} + \frac{QS^2}{B^2} = 1 \quad (G.32)$$

The equation of the inscribed circle gives :

$$\frac{O'Q^2}{A^2} + \frac{QK^2}{A^2} = 1 \quad (G.33)$$

Eliminating $\frac{O'Q^2}{A^2}$ between (G.32) and (G.33) results in

$$\frac{QS^2}{B^2} = \frac{QK^2}{A^2} \quad (G.34)$$

which is the same as (G.31). This gives :

$$r \sin v_k = A \sin E_k \times \frac{B}{A} = B \sin E_k \quad (G.35)$$

But

$$B = A (1 - e^2)^{\frac{1}{2}} \quad (G.36)$$

giving

$$r \sin v_k = A (1 - e^2)^{\frac{1}{2}} \sin E_k \quad (G.37)$$

Squaring gives

$$r^2 \sin^2 v_k = A^2 (1 - e^2) \sin^2 E_k \quad (G.38)$$

Squaring (G.30) gives

$$r^2 \cos^2 v_k = A^2 \cos^2 E_k - 2A^2 e \cos E_k + A^2 e^2 \quad (G.39)$$

Adding (G.38) and (G.39) gives

$$\begin{aligned} r^2 (\cos^2 v_k + \sin^2 v_k) &= A^2 [(\cos^2 E_k + \sin^2 E_k) - 2e \cos E_k \\ &\quad + e^2 (1 - \sin^2 E_k)] \end{aligned} \quad (G.40)$$

Taking the square root of (G.40) gives

$$\boxed{r = A(1 - e \cos E_k)} \quad (G.17)$$

To establish the relationship between true and eccentric anomalies, the geocentric radius in equation (G.30) is substituted by the expression given in (G.17). This gives

$$A(1 - e \cos E_k) \cos v_k = A \cos E_k - Ae \quad (G.41)$$

which results in

$$\boxed{\cos v_k = \frac{\cos E_k - e}{1 - e \cos E_k}} \quad (G.12)$$

Equation (G.13) is easily derived from (G.37), using (G.17), as

$$A(1 - e \cos E_k) \sin v_k = A(1 - e^2)^{\frac{1}{2}} \sin E_k \quad (G.42)$$

which gives:

$$\sin v_k = \frac{(1 - e^2)^{\frac{1}{2}} \sin E_k}{1 - e \cos E_k} \quad (\text{G.13})$$

REFERENCES AND BIBLIOGRAPHY

REFERENCES AND BIBLIOGRAPHY

- Agrotis, L. G., 1981. *Interpretation of Coordinate Differences and Refraction Modelling in Surveying Networks*. B.Sc. Thesis, Nottingham.
- Agrotis, L. G., Moore, T. and Ashkenazi, V., 1983. Determination of Satellite Orbits and the Analysis of LAGEOS Laser Range Observations. *Pres. Research Seminar on Space and Terrestrial Geodesy*. University of Nottingham.
- Anderle, R. J. 1974. Transformation of Terrestrial Survey Data to Doppler Satellite Datum. *Journal of Geophysical Research*, Vol.79, No.35.
- Anderle, R. J. 1982. Doppler Test Results of Experimental GPS Receiver. *NSWC TR 82-01*, Dahlgren, Virginia.
- Anderle, R. J. and Evans, A. G., 1982. Relative Positioning Test Using the Global Positioning System and Doppler Techniques. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.
- Ashkenazi, V., 1967. Solution and Error Analysis of Large Geodetic Networks. Part I. Direct Methods. *Survey Review*, NOS 146 and 147.
- Ashkenazi, V., 1969. Solution and Error Analysis of Large Geodetic Networks. Part II. Indirect Methods. *Survey Review*, NOS 151 and 152.
- Ashkenazi, V., 1970. Adjustment of Control Networks for Precise Engineering Surveys. *Chartered Surveyor*, January 1970.
- Ashkenazi, V., Crane, S. A. and Williams, J. W., 1979. The 1980 Redefinition of the OSGB Scientific Primary Triangulation Network. *Pres. XVII General Assembly of the IUGG*, Canberra.
- Ashkenazi, V., Crane, S. A. and Williams, J. W., 1981. The OSGB Scientific Network : combination of satellite and terrestrial data. *Ann. Geoph.*, Tome 37, No.1.
- 1 ✓ Ashkenazi, V., 1982. Positioning by Satellites : Transit and GPS. *Pres. General Meeting of the Hydrographic Society*.
- 27 ✓ Ashkenazi, V. and Agrotis, L. G., 1984. Satellite Positioning Systems and Geodetic Applications. *Pres. 8th UK Geophysical Assembly*, University of Newcastle-upon-Tyne.

- Ashkenazi, V., Agrotis, L. G. and Moore, T., 1984. Determination of Satellite Orbits and the Analysis of LAGEOS Laser Range Data. *Pres. 8th UK Geophysical Assembly*, University of Newcastle-upon-Tyne.
- Balmino, G., Moynot, B., Reigber, C., Müller, H., Rizos, H. and Bosch, W., 1983. An Improved GRIM3 Earth Gravity Model (GRIM3B). *Pres. XVIII General Assembly of the IUGG*, Hamburg.
- Black, H. D. and Eisner, A., 1982. Correcting Satellite Doppler Data for Tropospheric Effects. *Pres. 3rd International Symposium on The Use of Artificial Satellites for Geodesy and Geodynamics*, Greece.
- Blair, B. E. (ed), 1974. Time and Frequency : Theory and Fundamentals. *National Bureau of Standards Monograph 140*, Boulder, Colorado.
- Bock, Y., Abbot, R. I., Counselman, C. C., Gourevitch, S. A., King, R. W. and Paradis, A. R., 1983. Geodetic Accuracy of the Macrometer Model V-1000. *Pres. XVIII General Assembly of the IUGG*, Hamburg.
- 5 x Bomford, G., 1980. *Geodesy*, (Fourth Edition) Clarendon Press, Oxford.
- Bossler, J. D., Goad, C. C., and Bender, P. L., 1980. Using the Global Positioning System (GPS) for Geodetic Surveying. *Bull. Geodesique*, Vol.54.
- Bossler, J. D., 1982. Status of the Development of Geodetic Global Positioning Receivers. *Pres. ASCE Convention*, New Orleans, USA.
- Bureau International de l'Heure, 1981. *Annual Report for 1980*, Paris.
- Bureau International de l'Heure, 1982. *Annual Report for 1981*, Paris.
- Cazenave, A., Daillet, S. and Lambeck, K., 1977. Tidal Studies from the Perturbations in Satellite Orbits. *Phil. Trans. R. Soc. London A.284*, 595 - 606.
- Cartwright, D. E. and Tayler, R. J., 1971. New Computations of the Tide-Generating Potential. *Geophys. J. R. Astr. Soc.*, Vol.23.
- Christodoulidis, D. C., Smith, D. E., Dunn, P. J., Torrence, M. H., Knode, S. and Anders, S., 1982. The SL5 Geodetic Parameter Recovery Solution. *Pres. Third International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics*, Greece.

- Christodoulidis, D. C. and Smith D. E., 1983. The Role of Satellite Laser Ranging through the 1990's. *NASA Technical Memorandum 85104*, Greenbelt, Maryland.
- CIRA, 1972. *COSPAR International Reference Atmosphere*. Akademie - Verlag, Berlin.
- Counselman, C. C. and Shapiro, I. I., 1979. Miniature Interferometer Terminals for Earth Surveying. *Bull. Geodesique*, Vol.53.
- 7X Counselman, C. C. and Gourevitch, S. A., 1981. Miniature Interferometer Terminals for Earth Surveying : Ambiguity and Multipath with Global Positioning System. *IEEE Trans. on Geoscience and Remote Sensing*, Vol.GE-19, No.4.
- Counselman, C. C. and Steinbrecher, D. H., 1982. The Macrometer : A Compact Radio Interferometry Terminal for Geodesy with GPS. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.
- Diederich, P., Laue, H. and Rosetti, C., 1981. NAVSAT : A Global Civil Navigation Satellite System. *Pres. NAV 84, Conference of the Royal Institute of Navigation*, London.
- Dow, J., and Klinkrad, H., 1982. Orbit Determination and Control for the ERS-1 Alimetry Mission. *Pres. 3rd International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics*, Greece.
- Evans, A. G., Hermann, B. R. and Fell, P. J., 1981. Global Positioning System Sensitivity Experiment. *NAVIGATION : Journal of The Institute of Navigation*, Vol.28, No.2.
- Evans, A. G. and Hermann, B. R., 1982. Long Baseline (Transoceanic) Surveying Using The Global Positioning System and its Broadcast Ephemerides. *Pres. Third International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics*, Greece.
- Eplanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac, 1961, Fourth Impression 1977. H.M.S.O., London.
- 8X Fell, P. J., 1980. *Geodetic Positioning Using a Global Positioning System of Satellites*. Ph.D Thesis, Ohio State University, Report No.299.
- Fell, P. J., 1980. A Comparative Analysis of GPS Range, Doppler and Interferometric Observations for Geodetic Positioning. *Bull. Geodesique*, Vol.54.

- Gold, R., 1967. Optimal Binary Sequences for Spread Spectrum Multiplexing. *IEEE Trans. on Info.Theory*, pp.619-621.
- Goad, C. C. and Remondi, B. E., 1983. Initial Relative Positioning Results Using the Global Positioning System. *Pres. XVIII General Assembly of the IUGG*, Hamburg.
- Gough, R. J., 1978. *Satellite Doppler : Geocentric Positioning of the Triangulation of Great Britain*. Ph.D. Thesis, Nottingham.
- Haworth, R. T., 1984. Potential United Kingdom Participation in Project WEGENER. *Pres. 8th UK Geophysical Assembly*, University of Newcastle-upon-Tyne.
- Henrici, P., 1962. *Discrete Variable Methods in Ordinary Differential Equations*. John Wiley & Sons Inc.
- Hermann, B. R., 1981. Formulation for the NAVSTAR Geodetic Receiver System (NGRS). *NSWC TR 80-348*, Dahlgren, Virginia.
- Hermann, B. R. and Evans, A. G., 1981. A Demonstration of Relative Positioning Using Conventional GPS Doppler Receivers. *IEEE Telecommunications Conference*, New Orleans, U.S.A.
- Hermann, B. R. and Evans, A. G., 1982. Simultaneous Observations of Global Positioning System (GPS) Satellites with Two NAVSTAR Geodetic Receiver Systems (NGRS). *NGWC TR 81-472*, Dahlgren, Virginia.
- Hopfield, H. S., 1969. Two-Quartic Tropospheric Refractivity Profile for Correcting Satellite Data. *Journal of Geophysical Research*, Vol.74, No.18.
- Hopfield, H. S., 1972. Tropospheric Refraction Effects on Satellite Range Measurements. *APL/JHU Technical Digest*, Vol.11, No 4.
- Hopfield, H. S., 1977. Tropospheric Correction of Electromagnetic Ranging Signals to a Satellite : Study of Parameters. *IAG Symposium on Electromagnetic Distance Measurement and the Influence of Atmospheric Refraction*, Wageningen, The Netherlands.
- Hoskins, G. W., and Danchick, R. J., 1984. Joint Paper on Navy Navigation Satellite System Status and Future. *Pres. NAV 84, Conference of the Royal Institute of Navigation*, London.

Hothem, L. D. and Fronczek, C. J., 1983. Report on Test and Demonstration of MacrometerTM Model V-1000 Interferometric Surveyor. *FGCC Report : FGCC-IS-83-2*, Rockville, Maryland.

16 X ICD - GPS - 200, 1981. *NAVSTAR GPS Space Segment / Navigation User Interfaces*. Rockwell International, California.

17 X Janiczec, P. M. (ed), 1980. *NAVIGATION, Journal of the Institute of Navigation*. GPS Special Issue.

John Hopkins University and Stanford University, 1974.
A Satellite Freed of All But Gravitational Forces : 'TRIAD I'. *Pres. International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics*, Athens, Greece.

Kaplan, G. H. (ed), 1981. *United States Naval Observatory Circular No. 163*.

Krakiwsky, E. J. and Wells, D. E., 1971. *Coordinate Systems in Geodesy*. Department of Surveying Engineering, University of New Brunswick, Fredericton.

Kumar, M., 1982. An Unbiased Analysis of Doppler Coordinate Systems. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.

Lachapelle, G. and Beck, N., 1982. NAVSTAR / GPS Single Point Positioning at Sheltech Canada - Preliminary Results. *The Canadian Surveyor*, Vol. 36, No 1.

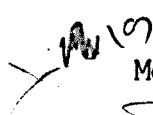
Lachapelle, G., Beck, N. and Heroux, P., 1982. NAVSTAR GPS Single Point Positioning Using Pseudo-Range and Doppler Observations. *Pres 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.

Lerch, F. J., Klosko, S. M., Laubscher, R. E. and Wagner, C. A., 1979. Gravity Model Improvement Using GEOS 3 (GEM 9 and GEM 10). *Journal of Geophysical Research*, Vol. 84, No 88.

Lerch, F. J., Klosko, S. M. and Patel, G. B., 1983. A Refined Gravity Model from Lageos (GEM-L2). *NASA Technical Memorandum 84986*, Greenbelt, Maryland.

Leroy, C. F., 1982. The Impact of GRS 80 on DMA Products. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.

Lucas, J. R., 1974. Radiation Pressure on a Spheroidal Satellite. *NOAA Technical Report NOS 61*, Rockville, MD.

- MacDoran, P. F., 1979. Satellite Emission Radio Interferometric Earth Surveying - SERIES - GPS Geodetic System. *Bull. Geodesique*, Vol.53.
- MacDoran, P. F., Spitzmesser, D. J. and Buennagel, L. A., 1982. SERIES : Satellite Emission Range Inferred Earth Surveying. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.
- Macomber, M. M., 1983. Personal Correspondence.
- Marini, J. W. and Murray, C. W., 1973. Correction of Laser Range Tracking Data for Atmospheric Refraction at Elevations Above 10 Degrees. *Goddard Space Flight Centre V-591-73-351*, Greenbelt, Maryland.
- Malyevac, C. and Anderle, R. J., 1982. Force model Improvement for NOVA Satellite. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.
- Martin, T. V., Eddy, W. F., Brenner, A., Rosen, B. and McCarthy, J., 1980. *Geodesyn System Description*, Vol.1. NASA Goddard Space Flight Center, Greenbelt, Maryland.
- McCarthy, D. D., 1982. The Use of Doppler Satellite Data for the Determination of Earth Orientation. *Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning*, Las Cruces.
-  McIntock, D. N., 1980. *Very Long Baseline Interferometry and Geodetic Applications*. Ph.D.Thesis, Nottingham.
- Melbourne, W. G., (ed), 1983. *MERIT Standards*.
- Melbourne, W. G., 1983a. GPS Precision Geodetic Systems. *Pres. XVIII General Assembly of the IUGG, Hamburg*.
- Miliken, R. J. and Zoller, C. J., 1978. Principle of Operation of NAVSTAR and System Characteristics. *NAVIGATION : Journal of The Institute of Navigation*, Vol.25, No.2.
- Moritz, H., 1984. Past and Future in Geodesy. *AVN*, Vol.1. Herbert Wichmann Verlag.
- Mueller, I. I., 1983. Report of the IAG / IAU Joint Working Group on the Establishment and Maintenance of a Conventional Terrestrial Reference System (COTES). *Pres. XVIII General Assembly of the IUGG, Hamburg*.

- NASA, 1981. NASA Geodynamics Program : Annual Report for 1981.
NASA Technical Memorandum 85126. Washington DC.
- O'Toole, J. W., 1976. CELEST Computer Program for Computing
Satellite Orbits. *NSWC/DL TR-3565, Dahlgren, Virginia.*
- Parkinson, B. W. and Gilbert, S. W., 1983. NAVSTAR : Global
Positioning System - Ten Years Later.
Proceedings of the IEEE, Vol.71, No 10.
- Payne, C. R., 1982. NAVSTAR Global Positioning System : 1982.
*Pres. 3rd International Geodetic Symposium on Satellite
Doppler Positioning, Las Cruces.*
- Porter, J., Kruh, P. and Sprosen, B., 1984. GPS NAVSTAR System
Overview. *Pres. NAV 84, Conference of the Royal
Institute of Navigation, London.*
- Prior, E. J., 1972. *Observed Effects of Earth-Reflected Radiation
and Hydrogen Drag on the Orbital Accelerations of Balloon
Satellites.* Geophysical Monograph 15, American Geophysical
Union, Washington DC.
- Reigber, C., Drewes, H., Müller, H., and Rizos, C., 1982.
Capabilities of LAGEOS Laser Ranging Data for Baseline
Determinations. *Pres. 3rd International Symposium on The
Use of Artificial Satellites for Geodesy and Geodynamics,
Greece.*
- Schwiderski, E. W., 1980. On Charting Global Ocean Tides.
Review of Geophysics and Space Physics 18 (1), 243 - 268.
- Sharman, P., 1982. The UK Satellite Laser Ranging Facility.
*SLR Technical Note No 1, Royal Greenwich Observatory,
Hertsmonceux, U.K.*
- Simms, M. L., 1982. GPS Geodetic Receiver System.
*Pres. 3rd International Geodetic Symposium on Satellite
Doppler Positioning, Las Cruces.*
- Spencer, A.J.M., et al, 1977. *Engineering Mathematics, Vol.2.*
Van Nostrand Reinhold.
- 25
Spilker, J. J., 1978. GPS Signal Structure and Performance
Characteristics. *NAVIGATION : Journal of the Institute
of Navigation, Vol.25, No 2.*
- 26
Sykes, R. M., 1979. *Translocation and Orbit Relaxation Techniques
in Satellite Doppler Tracking.* Ph.D. Thesis, Nottingham.

- Varnum, F. and Chaffee, J., 1982. Data Processing at the Global Positioning System Master Control Station.
Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning, Las Cruces.
- Wahr, J. M., 1979. *The Tidal Motions of a Rotating, Elliptical, Elastic and Oceanless Earth*. Ph.D. Thesis, University Of Colorado, Boulder, Colorado.
- Wahr, J. M., 1981a. Body Tides on an Elliptical, Rotating, Elastic and Oceanless Earth. *Geophys. J. R. Astr.Soc.*, Vol.64.
- Wahr, J. M., 1981b. The Forced Nutations of an Elliptical, Rotating, Elastic and Oceanless Earth.
Geophys. J.R. Astr.Soc., Vol.64.
- Ward, P., 1982. An Advanced NAVSTAR GPS Geodetic Receiver.
Pres. 3rd International Geodetic Symposium on Satellite Doppler Positioning, Las Cruces.
- Wilkins, G. A. (ed)., 1980. *A Review of The Techniques to be Used during Project MERIT To Monitor the Rotation of the Earth*. Published jointly by Royal Greenwich Observatory, Hertsmonceux, U.K. and Institut für Angewandte Geodäsie, Frankfurt, Germany.
- Wilkins, G. A. and Feissel, M., (eds), 1982. *Project MERIT: Report on the Short Campaign and Grasse Workshop with Observations and Results on Earth-Rotation during 1980 August-October*. Royal Greenwich Observatory, Hertsmonceux, U.K.
- Woolard, E. W., 1959. Inequalities in Mean Solar Time from Tidal Variations in the Rotation of the Earth.
The Astronomical Journal, 64, N^o 1269.